

# WEST Search History

DATE: Monday, March 03, 2003

**Set Name** **Query**  
side by side

**Hit Count** **Set Name**  
result set

*DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=ADJ*

L36	L33 and l14	1	L36
L35	L33 and l13	1	L35
L34	L33 and l12	0	L34
L33	L32 and l17	3	L33
L32	sodium-potassium adenosine triphosphorylase	3	L32
L31	L17 and l4	2	L31

*DB=USPT; PLUR=YES; OP=ADJ*

L30	l17 and l4	1	L30
L29	compound or composition or agent	948383	L29

*DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=ADJ*

L28	L4 and l20	1	L28
L27	L4 and l19	2	L27
L26	L4 and l18	0	L26
L25	L4 and l5	1	L25
L24	l5 same l4	0	L24
L23	l19 and l10	1	L23
L22	l20 and l10	1	L22
L21	l20 and l10	1	L21
L20	L13 with l9	46	L20
L19	L12 with l9	5	L19
L18	l14 with l9	38	L18
L17	l15 adj L16	824	L17
L16	transgenic	26173	L16
L15	rat or mouse or rodent	229213	L15
L14	L11 adj l3	13427	L14
L13	L11 adj l2	29140	L13
L12	L11 adj l1	27712	L12
L11	method	5683044	L11
L10	Dahl Salt sensitive rat	11	L10
L9	hypertension	35041	L9
L8	blood pressure or life span or renal pathology	52471	L8
L7	l4 with l5	0	L7
L6	L5 adj gene	2	L6





11



L5	hypertension susceptibility	10	L5
L4	Na,K-ATPase	83	L4
L3	screening	148614	L3
L2	identifying	267653	L2
L1	assaying or testing	463978	L1

END OF SEARCH HISTORY





1. 2.



10/040-722-7 01/4394:0  
attached  
STIC-Biotech/ChemLib

From: Whiteman, Brian  
Sent: Monday, March 03, 2003 2:41 PM  
To: STIC-Biotech/ChemLib  
Subject: seq search

CRFC  
10/040,722 filed 1/7/02 CON of 09/653,030 9/1/00  
Ruiz-Opazo

search seq id no: 3 against public DNA databases.

Thanks,  
Brian Whiteman, 11e12  
Patent Examiner - Art Unit 1635  
United States Patent and Trademark Office  
Crystal Mall 1, 11A16  
(703) 305-0775

Point of Contact:  
Beverly Shears  
Technical Info. Specialist  
CM1 1E05 Tel: 308-4994

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

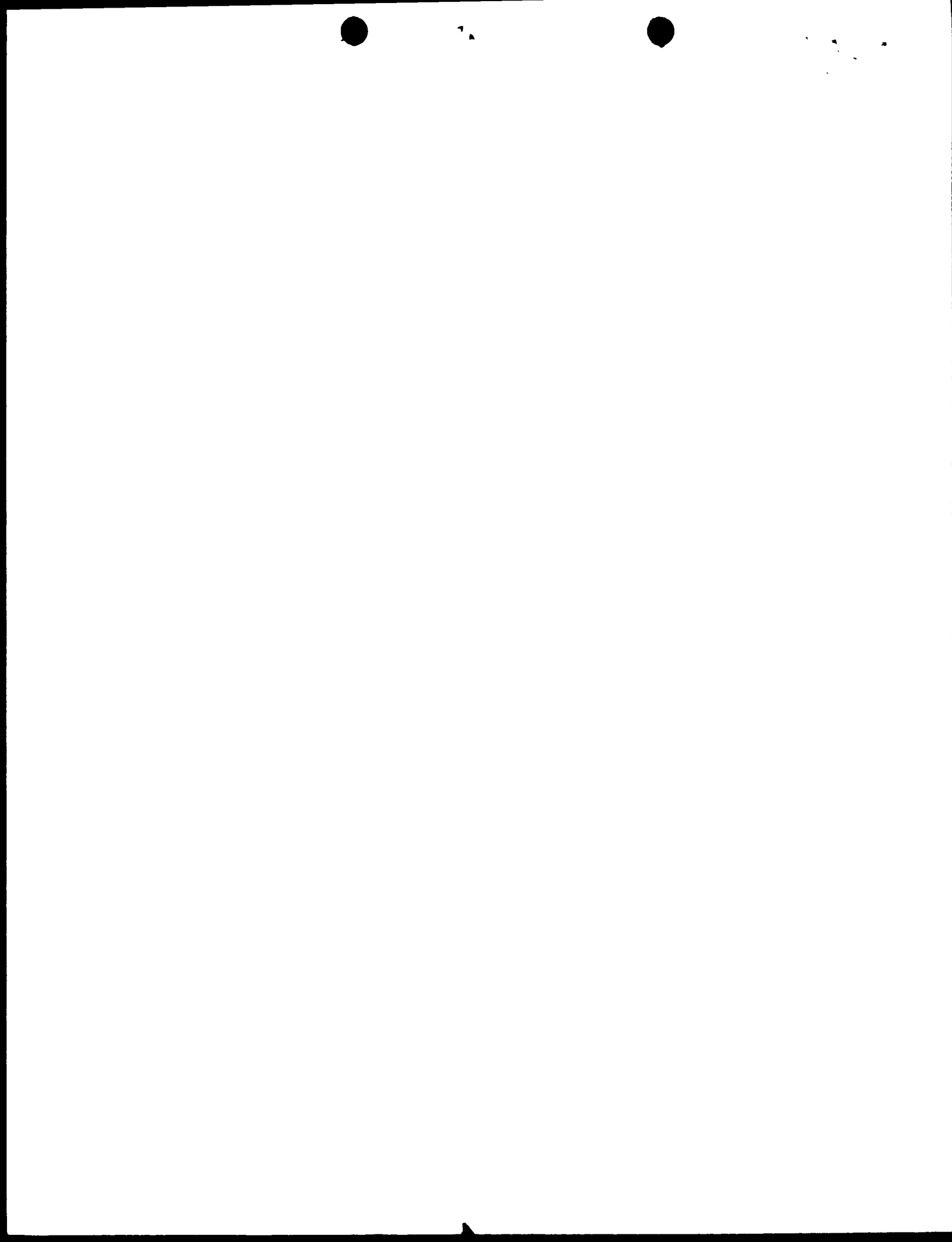
TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_







# SEARCH REQUEST FORM

Requestor's

Name: \_\_\_\_\_

Serial

Number: \_\_\_\_\_

Date: \_\_\_\_\_

Phone: \_\_\_\_\_

Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 03-11-03

Searcher: ~~100~~ Beverage 40499

Terminal time: 20

Elapsed time: \_\_\_\_\_

CPU time: \_\_\_\_\_

Total time: 23

Number of Searches: \_\_\_\_\_

Number of Databases: 1

### Search Site

\_\_\_\_\_ STIC

\_\_\_\_\_ CM-1

\_\_\_\_\_ Pre-S

### Type of Search

\_\_\_\_\_ N.A. Sequence

\_\_\_\_\_ A.A. Sequence

\_\_\_\_\_ Structure

\_\_\_\_\_ Bibliographic

### Vendors

\_\_\_\_\_ IG Suite

\_\_\_\_\_ STN

\_\_\_\_\_ Dialog

\_\_\_\_\_ APS

\_\_\_\_\_ Geninfo

\_\_\_\_\_ SDC

\_\_\_\_\_ DARC/Questel

☒ Other CGN







NH, Ternan  
 10/04-0722 Page 1  
 304 10 3  
 804

Description

AR175898 Se  
M14511 Pat

AR175895 Se

AR175897 Se  
D10359 Rat

X05882 Pat  
BC025811 M

BC021496 M

BC032187 N

BC025627 N  
M28647 Rat

BC010319 N  
BC003077 Hq

X04297	Hum
D00099	Hom

X02813 Shee

X03938 Pig  
M38445 Pig

L42173 Can.  
BC025037 M

J03230 Chi  
Z11798 B-m

U49238	Xenopus
U49239	Xenopus

X76108 A.a.

U82549 111  
AF286372 D

AY028630 D  
X05883 Ra

BC034645  
M14513 Ra

BC009394 H  
BC015566 H

BC009282 H:

AB018321 F  
AY008376 D

AY028629 D  
AL831997 H

BC020177  
AI.831991 H

AY008374 D

W. L. 431, Pa  
X. 58629 C. C

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

near PAI 2

1







```

Db 2041 GCTGTGGGCAAAATGCCGACGGCTGGAGATTAAAGTCATCATGGTCACAGAGACCAATCCA 2100
QY 2101 ATCAAGCCAAAGCCATTGGCTTAAGGGGGTGGGCAATTATCTCAGAAAGTAAAGACCCGTG 2160
Db 2101 ATCAAGCCAAAGCCATTGGCTTAAGGGGGTGGGCAATTATCTCAGAAAGTAAAGACCCGTG 2160
QY 2161 GAAGACATTTGGCTGCCCCGCTCAACATTCAGTGAACAGGTGAACCCCAAGATGCCAAG 2220
Db 2161 GAAGACATTTGGCTGCCCCGCTCAACATTCAGTGAACAGGTGAACCCCAAGATGCCAAG 2220
QY 2221 GCTGTGTAGTACATGAGAGTACTTGAAGACATGACCTCTGAGAGACTGGATGACATT 2280
Db 2221 GCTGTGTAGTACATGAGAGTACTTGAAGACATGACCTCTGAGAGACTGGATGACATT 2280
QY 2281 TTGGCGGTACCAACAGAGATTTGTCTTTGCTAGACCTCTCTCAACAGAGTCTCATTT 2340
Db 2281 TTGGCGGTACCAACAGAGATTTGTCTTTGCTAGACCTCTCTCAACAGAGTCTCATTT 2340
QY 2341 GTGAGAGGCTGCCAGCGGAGGGTGCATCGTGGCTGTCAAGGGGATGGTGCATGAC 2400
Db 2341 GTGAGAGGCTGCCAGCGGAGGGTGCATCGTGGCTGTCAAGGGGATGGTGCATGAC 2400
QY 2401 TCTCCAGCTTTGAAAAAGGAGATTTGGGGTGGCCATGGGGATTTGGCTGGATGTG 2460
Db 2401 TCTCCAGCTTTGAAAAAGGAGATTTGGGGTGGCCATGGGGATTTGGCTGGATGTG 2460
QY 2461 TCCAGCAGCTCTGACATGATTTCTTGTGGATGACAACTTTGCTCCATCGTGCATGGA 2520
Db 2461 TCCAGCAGCTCTGACATGATTTCTTGTGGATGACAACTTTGCTCCATCGTGCATGGA 2520
QY 2521 GTGAGAGAGGTGCTGATATTTGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2580
Db 2521 GTGAGAGAGGTGCTGATATTTGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2580
QY 2581 AGTAACTTTGGGAAATATACCCCTCTTATATTTATTTATTTATTTATTTATTTATTTATTT 2640
Db 2581 AGTAACTTTGGGAAATATACCCCTCTTATATTTATTTATTTATTTATTTATTTATTTATTT 2640
QY 2641 CTGGGACCGGTGACCATCTCTGATGATGACTTGGGACCTGACATGCTTCCCGCATCTCT 2700
Db 2641 CTGGGACCGGTGACCATCTCTGATGATGACTTGGGACCTGACATGCTTCCCGCATCTCT 2700
QY 2701 CTGGCTATTAAGAGGCTGAAGATGACATCATGAAAGAGGAGCCAGAAATCCCAAAACG 2760
Db 2701 CTGGCTATTAAGAGGCTGAAGATGACATCATGAAAGAGGAGCCAGAAATCCCAAAACG 2760
QY 2761 GACAAACTTGTGAACGAGGCTGTGATGACATGAGGCTTATGACAGATGGTATGATCCAG 2820
Db 2761 GACAAACTTGTGAACGAGGCTGTGATGACATGAGGCTTATGACAGATGGTATGATCCAG 2820
QY 2821 GCTCTGGAGAGGCTTCTTCACTTATTTTGTGATTTGGCTGAGAACGGTTTCCGCTCTT 2880
Db 2821 GCTCTGGAGAGGCTTCTTCACTTATTTTGTGATTTGGCTGAGAACGGTTTCCGCTCTT 2880
QY 2881 CACCTTTGGGATCCGAGAGACCTGGGATGACCGCTGATCATGATGATGAGAGACAGC 2940
Db 2881 CACCTTTGGGATCCGAGAGACCTGGGATGACCGCTGATCATGATGATGAGAGACAGC 2940
QY 2941 TACGGGACAGAGTGAACCTTACGAGAGAGAGATTTGTGAGTTCACCTGCCACAGGCTC 3000
Db 2941 TACGGGACAGAGTGAACCTTACGAGAGAGAGATTTGTGAGTTCACCTGCCACAGGCTC 3000
QY 3001 TTTCTTGTAGTATCGTGTAGTGTGAGTGGCTGACTTGGTCACTCTGCAAGACCAAGAG 3060
Db 3001 TTTCTTGTAGTATCGTGTAGTGTGAGTGGCTGACTTGGTCACTCTGCAAGACCAAGAG 3060
QY 3061 AATTCTGTCTTCCAGGAGGAAATGAACAAGATCTTAATTTATTTATTTATTTATTTATTTAG 3120
Db 3061 AATTCTGTCTTCCAGGAGGAAATGAACAAGATCTTAATTTATTTATTTATTTATTTATTTAG 3120
QY 3121 ACAGCTTGTGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3180

```

```

Db 3121 ACAGCTTGTGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3180
QY 3181 CCCCTCAAACTTACTTGTGATGATTTGTGCTTTCCCTTACTCCCTTCTCATCTTCTGTAT 3240
Db 3181 CCCCTCAAACTTACTTGTGATGATTTGTGCTTTCCCTTACTCCCTTCTCATCTTCTGTAT 3240
QY 3241 GACGAGGTGGAGAGTCTATCATCAGGAGAGGAGCCCTGCGGCTGGGTGAGAGAAAGAAC 3300
Db 3241 GACGAGGTGGAGAGTCTATCATCAGGAGAGGAGCCCTGCGGCTGGGTGAGAGAAAGAAC 3300
QY 3301 TACTACTAGCCCACTGCTCCCTGACAGCGTCTGGAATGAGATGATGATGATGATGATGATGAT 3360
Db 3301 TACTACTAGCCCACTGCTCCCTGACAGCGTCTGGAATGAGATGATGATGATGATGATGATGAT 3360
QY 3361 TACCCCCCTTTTGTATCTTCAAGTCTTGGAGCTGGAGACTTCACTTCTGATGAGAAAGCA 3420
Db 3361 TACCCCCCTTTTGTATCTTCAAGTCTTGGAGCTTGGAGACTTCACTTCTGATGAGAAAGCA 3420
QY 3421 CCAGAGCATGGGAGATCCAGACGCTCTGGAATGAGATGATGATGATGATGATGATGATGATGAT 3480
Db 3421 CCAGAGCATGGGAGATCCAGACGCTCTGGAATGAGATGATGATGATGATGATGATGATGATGAT 3480
QY 3481 GGGAGAGGCTCCCGAAAAACACCGTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
Db 3481 GGGAGAGGCTCCCGAAAAACACCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
QY 3541 CTTTCTTTTGTGTAAGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Db 3541 CTTTCTTTTGTGTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
QY 3601 GGATTTTTCATTAAGATGAGTATTAAGAGAA 3636
Db 3601 GGATTTTTCATTAAGATGAGTATTAAGAGAA 3636

```

## RESULT 2

RATATP1 3636 bp mRNA linear PDB 27-APP-1993  
LOCUS Rat Na<sup>+</sup>,K<sup>+</sup>-ATPase alpha isoform catalytic subunit mRNA, complete

## DEFINITION

Accession M14511.1  
Version M14511.1  
Keywords Rat male brain (strain CD) cDNA to mRNA, clone B1-4  
Source Rattus norvegicus  
Organism Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

## REFERENCE

1 (bases 1 to 3636)  
Shull, G.E., Greeb, J. and Lingrel, J.B.  
Molecular cloning of three distinct forms of the Na<sup>+</sup>,K<sup>+</sup>-ATPase  
alpha-subunit from rat brain  
Biochemistry 25 (25), 8125-8132 (1986)

## JOURNAL

87128908  
MEDLINE  
PUBMED  
COMMENT  
by G.E.Shull, 18-FEB-1987.  
Location/Qualifiers  
1. 3636  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/note="(Na<sup>+</sup> and K<sup>+</sup>) ATPase, alpha catalytic subunit  
precursor"

## FEATURES

1. 3636  
/db\_xref="taxon:10116"  
/note="(Na<sup>+</sup> and K<sup>+</sup>) ATPase, alpha catalytic subunit  
precursor"

## CDs

1. 3636  
/db\_xref="taxon:10116"  
/note="(Na<sup>+</sup> and K<sup>+</sup>) ATPase, alpha catalytic subunit  
precursor"

## CDs

1. 3636  
/db\_xref="taxon:10116"  
/note="(Na<sup>+</sup> and K<sup>+</sup>) ATPase, alpha catalytic subunit  
precursor"

## CDs

1. 3636  
/db\_xref="taxon:10116"  
/note="(Na<sup>+</sup> and K<sup>+</sup>) ATPase, alpha catalytic subunit  
precursor"

## CDs

1. 3636  
/db\_xref="taxon:10116"  
/note="(Na<sup>+</sup> and K<sup>+</sup>) ATPase, alpha catalytic subunit  
precursor"

## CDs

1. 3636  
/db\_xref="taxon:10116"  
/note="(Na<sup>+</sup> and K<sup>+</sup>) ATPase, alpha catalytic subunit  
precursor"

## CDs

1. 3636  
/db\_xref="taxon:10116"  
/note="(Na<sup>+</sup> and K<sup>+</sup>) ATPase, alpha catalytic subunit  
precursor"

## CDs

1. 3636  
/db\_xref="taxon:10116"  
/note="(Na<sup>+</sup> and K<sup>+</sup>) ATPase, alpha catalytic subunit  
precursor"







QY 1861 TACCTAGAGGCTGGGGGCTTGGAGAGCGTGTGTAGATTCTTCCACCTCTCTTCTGCT 1920  
 Db 1861 TACCTAGAGGCTGGGGGCTTGGAGAGCGTGTGTAGATTCTTCCACCTCTCTTCTGCT 1920  
 QY 1921 GACGACAGTTTCCCGAAGGCTTCCAGTTTGAACCTGATGAAGCAATTTTCCCGTGGAT 1980  
 Db 1921 GACGACAGTTTCCCGAAGGCTTCCAGTTTGAACCTGATGAAGCAATTTTCCCGTGGAT 1980  
 QY 1981 AACCTCTGCTGCTGGGTCTTATCTCATGATTGACCCCTCTGAGCTGCTGTCCCGAT 2040  
 Db 1981 AACCTCTGCTGCTGGGTCTTATCTCATGATTGACCCCTCTGAGCTGCTGTCCCGAT 2040  
 QY 2041 GCTGTGGCAAAATGCCCGAGCGCTGGATTAAGGTCATCATGGTCAAGAGACCATCA 2100  
 Db 2041 GCTGTGGCAAAATGCCCGAGCGCTGGATTAAGGTCATCATGGTCAAGAGACCATCA 2100  
 QY 2101 ATCAGAGCCAAAGGCTATTGAGGGGGTGGGCAATTATCTGAGAGTTACAGACCGTG 2160  
 Db 2101 ATCAGAGCCAAAGGCTATTGAGGGGGTGGGCAATTATCTGAGAGTTACAGACCGTG 2160  
 QY 2161 GAAGACATTGCTGCCCGCTCAAACTTCAGAGAACAGAGTGAACCCCGAGATGCCAAG 2220  
 Db 2161 GAAGACATTGCTGCCCGCTCAAACTTCAGAGAACAGAGTGAACCCCGAGATGCCAAG 2220  
 QY 2221 GCTGTGTAGTATGATGAGCACTGATGAAGAGCATGACCTTGAAGAGCTGGATGAT 2280  
 Db 2221 GCTGTGTAGTATGATGAGCACTGATGAAGAGCATGACCTTGAAGAGCTGGATGAT 2280  
 QY 2281 TTGCGGTACCAACGAGATTTGCTTGTAGAGCTCTCTCCAGAGAGTCAATTCAT 2340  
 Db 2281 TTGCGGTACCAACGAGATTTGCTTGTAGAGCTCTCTCCAGAGAGTCAATTCAT 2340  
 QY 2341 GTGAGAGGCTGCCAGCGGAGAGGTCATCGTGGCTGTCAAGGGGATGGTGTCAATGAC 2400  
 Db 2341 GTGAGAGGCTGCCAGCGGAGAGGTCATCGTGGCTGTCAAGGGGATGGTGTCAATGAC 2400  
 QY 2401 TCTCAGCTTTGAAAAAGCAGATTTGGGGTTGCCATGGGGANTGTGGTGGATGTG 2460  
 Db 2401 TCTCAGCTTTGAAAAAGCAGATTTGGGGTTGCCATGGGGANTGTGGTGGATGTG 2460  
 QY 2461 TCTCAGCTTTGAAAAAGCAGATTTGGGGTTGCCATGGGGANTGTGGTGGATGTG 2520  
 Db 2461 TCTCAGCTTTGAAAAAGCAGATTTGGGGTTGCCATGGGGANTGTGGTGGATGTG 2520  
 QY 2521 GTAGAGAGGTCGTCTGATATTGATTAAGTGAAGAAATCCATTGCTTACACCTTAA 2580  
 Db 2521 GTAGAGAGGTCGTCTGATATTGATTAAGTGAAGAAATCCATTGCTTACACCTTAA 2580  
 QY 2581 AGTAACATTCCGGAATCAACCCCTTCTGATATTATTATTGCAACATTCACCTGCC 2640  
 Db 2581 AGTAACATTCCGGAATCAACCCCTTCTGATATTATTATTGCAACATTCACCTGCC 2640  
 QY 2641 CTGGGACCTGACCATCTCTGATTAAGTGAAGAAATCCATTGCTTACACCTTAA 2700  
 Db 2641 CTGGGACCTGACCATCTCTGATTAAGTGAAGAAATCCATTGCTTACACCTTAA 2700  
 QY 2701 CTGGGACCTGACCATCTCTGATTAAGTGAAGAAATCCATTGCTTACACCTTAA 2760  
 Db 2701 CTGGGACCTGACCATCTCTGATTAAGTGAAGAAATCCATTGCTTACACCTTAA 2760  
 QY 2761 GACAAACCTGTGAACGAGCTGATCAAGTGAAGAAATCCATTGCTTACACCTTAA 2820  
 Db 2761 GACAAACCTGTGAACGAGCTGATCAAGTGAAGAAATCCATTGCTTACACCTTAA 2820  
 QY 2821 GACCTGTGGGATCCGAGAGCTGGGATGACCGGATGATTAAGTGAAGAAATCC 2880  
 Db 2821 GACCTGTGGGATCCGAGAGCTGGGATGACCGGATGATTAAGTGAAGAAATCC 2880  
 QY 2881 GACCTGTGGGATCCGAGAGCTGGGATGACCGGATGATTAAGTGAAGAAATCC 2940  
 Db 2881 GACCTGTGGGATCCGAGAGCTGGGATGACCGGATGATTAAGTGAAGAAATCC 2940  
 QY 2941 TACGGGACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3000

Db 2941 TACGGGACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3000  
 QY 3001 TACCTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
 Db 3001 TACCTGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
 QY 3061 AATTGTCTTCCAGAGGAGTGAAGAAATCCATTGCTTACACCTTAA 3120  
 Db 3061 AATTGTCTTCCAGAGGAGTGAAGAAATCCATTGCTTACACCTTAA 3120  
 QY 3121 AAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
 Db 3121 AAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
 QY 3181 CCCCTCAACCTACTTGT 3240  
 Db 3181 CCCCTCAACCTACTTGT 3240  
 QY 3241 GAGAGAGTGGGAGAGTCAATCATGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
 Db 3241 GAGAGAGTGGGAGAGTCAATCATGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
 QY 3301 TACTACTAGCCCACTGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
 Db 3301 TACTACTAGCCCACTGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
 QY 3361 TACTACTAGCCCACTGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420  
 Db 3361 TACTACTAGCCCACTGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420  
 QY 3421 CCMAAGATGTGGGATCCAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480  
 Db 3421 CCMAAGATGTGGGATCCAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480  
 QY 3481 GGGAGAGGCTGCCCGGAG 3540  
 Db 3481 GGGAGAGGCTGCCCGGAG 3540  
 QY 3541 CCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 3600  
 Db 3541 CCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 3600  
 QY 3601 GGAATTTTCAAAATAAGATGGCTATTATACGAA 3636  
 Db 3601 GGAATTTTCAAAATAAGATGGCTATTATACGAA 3636  
 RESULT 3  
 ARI75895 3636 bp DNA linear PAT 17-DEC-2001  
 LOCUS ARI75895  
 DEFINITION Sequence 5 from patent US 6309874.  
 ACCESSION ARI75895  
 VERSION ARI75895.1 GI:17917194  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 3636)  
 AUTHORS Belusa, R.  
 TITLE Selection marker  
 JOURNAL Patent: US 6309874-A 5 30-OCT-2001;  
 FEATURES  
 source 1..3636  
 BASE COUNT 850 a 967 c 995 g 824 t  
 ORIGIN  
 Query Match 99.7%; Score 3626.4; DB 6; Length 3636;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 3630; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 GGAGCCTCGGGGAGAGAGCGGAGACGTTGAGCGGGGCGGAGCGGAGCAGACGAGCG 60







[illegible]

LOCUS	AR175896	3636 bp	DNA	linear	PAT 17-DEC-2001
DEFINITION	Sequence 6 from patent US 6309874.				
ACCESSION	AR175896				
VERSION	AR175896.1	GI:17917195			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 3636)				
TITLE	Belusa, R.				
JOURNAL	Selection marker				
FEATURES	Patent: US 6309874-A 6 30-OCT-2001;				
	Location/Qualifiers				
	1..3636				
	/organism="unknown"				
BASE COUNT	850 a	967 c	995 g	824 t	
ORIGIN					
Query Match	99.8%; Score 3626.4; DB 6; Length 3636;				
Best Local Similarity	99.8%; Pxd. No. 0;				
Matches 3630; Conservative	0; Mismatches 6; Indels 0; Gaps 0				
1	GGAGCCTTGGCGGAGGAGGCGGACACAGCTGGCGACGCGCGCGGCGGAGCGGCGGAGCAGCAGCG	60			
1	GGAGCCTTGGCGGAGGAGGCGGACACAGCTGGCGACGCGCGGCGGAGCGGAGCAGCAGCGG	60			
61	GGCGCAGCAGCGCGGCGCTCGGTCGGGGGCGCGCGGCGCTCTCTTTCTTCTTCGCGCG	120			
61	GGCGCAGCAGCGCGGCGCTCGGTCGGGGGCGCGCGGCGCTCTCTTTCTTCTTCGCGCG	120			
121	CAGCCCTTAGTTCCTCGGCTCGGTCGCCCGGCTGCCACTCTCCCGAGCGGGAGCTGCTCT	180			
121	CAGCCCTTAGTTCCTCGGCTCGGTCGCCCGGCTGCCACTCTCCCGAGCGGGAGCTGCTCT	180			
181	CTCCTCTTTAGTCTCCAGCCACAGACCCCGGCGCGGGGCGCGGAGCGCGGCTCACCATG	240			
181	CTCCTCTTTAGTCTCCAGCCACAGACCCCGGCGCGGGGCGCGGAGCGCGGCTCACCATG	240			
241	GGGAAGGGGGTTGACGACGACAGATAGCCCGACGCTGTATGAAACATG3G9ACAAG	300			
241	GGGAAGGGGGTTGACGACGACAGATAGCCCGACGCTGTATGAAACATG3G9ACAAG	300			
301	AAGACCAAGAGCGAAGAGGAAAGGACAT3ACGAACTCAAGAAAGGAAGTGCTATG	360			
301	AAGACCAAGAGCGAAGAGGAAAGGACAT3ACGAACTCAAGAAAGGAAGTGCTATG	360			







RESULT		PAT 17-DEC-2001
LOCUS	AR175897	linear
DEFINITION	AR175897 7 from patent US 6309874.	DNA
ACCESSION	AR175897	
VERSION	AR175897.1 GI:17917196	
KEYWORDS	.	
SOURCE	Unknown.	
ORGANISM	Unclassified.	
REFERENCE	1 (bases 1 to 3636)	
AUTHORS	Belusa,R.	
TITLE	Selection marker	
JOURNAL	Patent: US 6309874-A 7 30-OCT-2001;	
FEATURES	Location/Qualifiers 1..3636 /organism="unknown"	
BASE COUNT	850 a      966 c      995 g      825 t	
ORIGIN		

Query Match	99.6%	Score 3621.6	DB 6	Length 3636
Best Local Similarity	99.8%	Pred. No. 0	Indels 9	Gaps 0
Matches 3627	Conservative	0	Mismatches	
QY 1	GGAGCTCGGGGAGAGGACGACGATGAGCAGCGCGCGCGGACGCGGACAGAGCG	60		
DB 1	GGAGCTCGGGGAGAGGAGGCGGACGATGAGCAGCGCGCGGACGCGGACAGAGCG	60		
QY				
DB 61	GGCGCAGCAGCGCGCGCTCGTCCGGGGGCGCGCGCTCCCTCTTCTCGCGCG	120		
QY				
DB 61	GGCGCAGCAGCGCGCGCTCGTCCGGGGGCGCGCGCTCCCTCTTCTCGCGCG	120		
QY				
DB 121	CAGCCCTAGTTCGCGCTCTCGGCTCCCGCGCTCACTCTCCAGCGGAGCTGCT	180		
QY				
DB 121	CAGCCCTAGTTCGCGCTCTCGGCTCCCGCGCTCACTCTCCAGCGGAGCTGCT	180		
QY				
DB 181	CTCCTCTTCTAGTCTCCAGCCAGAGACCGCGCGCGCGCGCGCGCCACATG	240		
QY				
DB 181	CTCCTCTTCTAGTCTCCAGCCAGAGACCGCGCGCGCGCGCGCGCCACATG	240		
QY				
DB 241	GGGAGGGGGGTTGAGCAGACAGATATAGCCGAGCTGTATCAGAACATGGGAGCAG	300		
QY				
DB 241	GGGAGGGGGGTTGAGCAGACAGATATAGCCGAGCTGTATCAGAACATGGGAGCAG	300		
QY				
DB 301	AGAGCAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	360		
QY				
DB 301	AGAGCAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	360		
QY				
DB 361	GAGCAGCATTAATCTGAGCTGAGTGAATCTGTAATAGCGAAGAGATTGAGCGA	420		
QY				
DB 361	GAGCAGCATTAATCTGAGCTGAGTGAATCTGTAATAGCGAAGAGATTGAGCGA	420		
QY				
DB 421	GGCTTAACACCGCGAAGGGCGCTGAGTCTGAGCTCGGATGCGCCCAACGCTTCAG	480		
QY				
DB 421	GGCTTAACACCGCGAAGGGCGCTGAGTCTGAGCTCGGATGCGCCCAACGCTTCAG	480		
QY				
DB 481	CCGCTTCCACTACTCTCCGATGGGTCAATTCGTGCGAGCTGTTCGCTGCTTCC	540		
QY				
DB 481	CCGCTTCCACTACTCTCCGATGGGTCAATTCGTGCGAGCTGTTCGCTGCTTCC	540		
QY				
DB 541	ATGTACGTGAGATTGAGAGCATTTGTTCTTGCTTATGCGATCGAAGTCTACA	600		
QY				
DB 541	ATGTACGTGAGATTGAGAGCATTTGTTCTTGCTTATGCGATCGAAGTCTACA	600		
QY				
DB 601	GAGAGGAACCAACCAATGATCTGACTCGGGGCGTGTCTGTCTGTCATC	660		
QY				
DB 601	GAGAGGAACCAACCAATGATCTGACTCGGGGCGTGTCTGTCTGTCATC	660		
QY				
DB 661	ATTAATGAGGTCTTCTTATATACAGAGCAAAAGCTCCAGATATGSAATCTTC	720		
QY				
DB 661	ATTAATGAGGTCTTCTTATATACAGAGCAAAAGCTCCAGATATGSAATCTTC	720		
QY				
DB 720	ATTAATGAGGTCTTCTTATATACAGAGCAAAAGCTCCAGATATGSAATCTTC	720		







[illegible]

COMMENT	FEATURES	source
Found by the N-terminal amino acid sequence of the mature alpha-subunit additional amino acid residues were cleaved during the post-translational processing, and pRNP11 was found to carry a CDNA encoding the alpha subunit of the rat kidney-type.	Location/Qualifiers 1. .3407 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db_xref="taxon:10116" /clone="pRNP11" /tissue_type="brain" 15. .3086 /codon_start=1 /product="Na+,K+-ATPase alpha-subunit precursor" /protein_id="BA01189.1" /db_xref="GI:220824"	CDS
	sig_peptide mat_peptide 18. .29 30. .3083 /product="Na+,K+-ATPase alpha-subunit mature peptide"	
	BASE COUNT 822 a 881 c 913 g 791 t	ORIGIN
Query Match 92.1% Score 3347.6; DB 10; Length 3407; Best local similarity 99.3%; Pred. No. 0; Matches 3383; Conservative 0; Mismatches 19; Indels 4; Gaps 2;		
228 CGCCGCCACCATGCGGAAAGGCGGTTTGACGAGACAGATATGAGCCCGCAGCTGATACGA 287		
5 CGCGCGCCACCATGCGGAAAGGCGGTTTGACGAGACAGATATGAGCTGATATGAGA 64		
288 ACATATGGGAGACAAAGAGACCAAGAAAGGCGCAACAAGGAGACATATGACGAAATTCAGAA 347		
65 ACATATGGGAGACAAAGAGACCAAGAAAGGCGCAACAAGGAGACATATGACGAAATTCAGAA 124		
348 GGAAGTGTCTATGACGACGACATAAACTAGCTAGCTGGATATATCTCATCGTAAATACGGAAC 407		
125 GGAAGTGTCTATGACGACGACATAAACTAGCTAGCTGGATATATCTCATCGTAAATACGGAAC 184		
408 AGATTTGAGCCGAGGCTTAACAACCCGCAAGGCGCGTGGATCTCTGCTCGGGAATGCGCC 467		
185 AGATTTGAGCCGAGGCTTAACAACCCGCAAGGCGCGCTAGATCTCTGCTCGGGAATGCGCC 244		
468 CAAGCCCGCTCAAGCCCGCTCCGACACTACCCGAGTCCGATCCAAATTCGTGGGCGAGTGT 527		
245 CAAGCCCGCTCAAGCCCGCTCCGACACTACCCGAGTCCGATCCAAATTCGTGGGCGAGTGT 304		
528 CGATGGCTTTCATATGATGTACTGTGGATATGAGACCATTTCTTGTGTTCTTACCTTAATGCGCAT 587		
305 CGATGGCTTTCATATGATGTACTGTGGATATGAGACCATTTCTTGTGTTCTTACCTTAATGCGCAT 364		
588 CGAAGTGTCTATGAGAAAGGAAACCAACCAATATGATATTTGACTCGGAGGTGGGCTGCTC 647		
365 CGAAGTGTCTATGAGAAAGGAAACCAACCAATATGATATTTGACTCGGAGGTGGGCTGCTC 424		































[illegible]

QY	2941	TAAAGGAGCAGTGGACCTCAGACAAAGGAAATGTGAGTTCACCTGCAACAGCGC	3000
Db	2947	TACGGAGCAGCTGACCTCAGACAGAGAAATCGTGAGTTCACTCCATACAGCG	3006
QY	3001	TTCTTTGTCAATATCGTGATAGTGAGTGAGGCTGACTTTGTCATCTCGAAGACCAAGG	3060
Db	3007	TTCTTTGTCAATATTTGTGTAGTGCATGGGCGCACTTGTGCATCTCGAAGACCAAGG	3066
QY	3061	AATTCTGCTTCCGAGCGAGGAAATGAAGACAAGATCTTAATATTATTTGGCCCTTTGAAG	3120
Db	3067	AATTCTGCTTCCGAGCGAGGATGAAGACAAGATCTTGATATTGGCCCTTTGAAGAG	3126
QY	3121	ACAGCTCTGTGCTTCTCTGCTCTCTACTGCCCTTGAGATGGGTGACGCCCTTAGATGTAT	3180
Db	3127	ACAGCCCTGTGCTTCTTATCTCTACTGCCCCGGGAAATGGGAGCAGCCCTTAGATGTAT	3186
QY	3181	CCCCCTAAACCTACTTGGTGGTCTGTGTGCTTCCCCCACTCCCTTCATCTTGGTGAT	3240
Db	3187	CCCCCTAAACCTAATGGTGGTGTGTGTGCTTCCCTCACTCCCTTCATCTTGTGTAT	3246
QY	3241	GACGAGTGCAGAAAGCTCATCATGAGCGAGCCCTGCGGCTGATGATGAGAGAAACG	3300
Db	3247	GACGAGTGCAGAAAGCTCATTTATCAGCGAGCCCTGCGGCTGATGAGAGAAACG	3306
QY	3301	TACTACTAGCCCACTGCCCTGACGCGGTGGAATTTGTGCAACACTGACCTTACCCC	3360
Db	3307	TACTACTAGCCCACTGCCCTGACGCGGTGGAATTTGTGCAACACTGATCCATCCC	3366
QY	3361	TACCCCCCTTTGTGTACTTCAAGCTTGGAGCTCGGAACCTTACCCCTGTAGGAAACA	3420
Db	3367	TACCCCCCTTTGTGTACTTCAAGCTTGGAGCTCGGAACCTTACCCCTGTAGGAAACA	3426
QY	3421	CCAAAGCATGTGGGATCCAGACGTCCTCTGTAATGAAATGTATGTGTAATGGGAGGCG--	3478
Db	3427	CCAAAGCATGTGGGATCCAGACGTCCTCTGTAATGAAATGTATGTGTAATGGGAGGCG	3486
QY	3479	----GGGGGAGAGGCTGCGCCGAAAAACACCGCTGACGCGGAGCGACAGCGGCGAAGTTTA	3534
Db	3487	GGCGGGGGGAGAGGCTGCTGAGAAACACCGTGTACGGGAGCGACAGCGGCGAAGTTTA	3546
QY	3535	TATGTGCTTTTGTTTTGTAAAAAGAAAAAAGCTGGAAAAAGCTGAAAAGTTACGTTTT	3594
Db	3547	TACGTGCTTTTGTTTTGTAAAAAGAAAAAAGCTGGAAAAAGCTGAAAAGTTACGTTTT	3606
QY	3595	AATATCGATTTTTCAAATAAAGATGGCTATTATTAACGGAA	3636
Db	3607	AATATCGATTTTTCAAATAAAGATGGCTATTATTAACGGAA	3648

RESULT 10	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	AUTHORS	JOURNAL	PEPABP	COMMENT
BC025618	BC025618	Mus musculus, clone MGC:38134 IMAGE:5320976, mRNA, complete cds.	BC025618	BC025618.1	GI:19343735	MGC.	house mouse.	1 (bases 1 to 3686)	Submitted (06-MAR-2002)	Strausberg,R	National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	Contact: MGC help desk Email: <a href="mailto:cgaps-remail.nih.gov">cgaps-remail.nih.gov</a> Tissue Procurement: Jeffrey Green M.D. CDNA Library Preparation: Life Technologies, Inc.







QY 1300 AACTGCTGTGAAGAACTTGGAGCTGTGGAGACCTTGGGGTTCACATCCACCATCTTGC 1359  
| | | | |  
Db 1315 AACTGCTGTGAAGAACTTGGAGCTGTGGAGACCTTGGGGTTCACATCCACCATCTTGC 1374  
| | | | |  
QY 1360 TCCGACAAAGATGTGAACCTCTGACTCAGAACCGGATGACAGTGGCTCACAATGGTTTAC 1419  
| | | | |  
Db 1375 TCACACAAAGACGGGAATCTCTACTCAGAACCGGATGACAGTGGCTCACAATGGTTTAC 1434  
| | | | |  
QY 1420 AATCAATCCATGAAAGCTGACACCAAGAGAAATCAAGTGGGGTCTCTTTGACAAAGCG 1479  
| | | | |  
Db 1435 AATCAAGATCCATGAAAGCTGACACCAAGAGAAATCAAGTGGGGTCTCTTTGACAAAGCG 1494  
| | | | |  
QY 1480 TCACGACACCTGTGCTCTGTCCGAAATGTGGTGTCTGTAAACAGGCGAGTGTTCAG 1539  
| | | | |  
Db 1495 TCACGACACCTGTGCTCTGTCCGAAATGTGGTGTCTGTAAACAGGCGAGTGTTCAG 1554  
| | | | |  
QY 1540 GCTAACCCAGAAAACTTGCCTATCTTAAAGCGTGTGAGTGGCGGAGATGCTTCCGAGTGC 1599  
| | | | |  
Db 1555 GCTAACCCAGAAAACTTGCCTATCTTAAAGCGGCGAGTGGCGGAGATGCTTCCGAGTGC 1614  
| | | | |  
QY 1600 GCGCTCTTAAAGTGTATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1659  
| | | | |  
Db 1615 GCGCTCTTAAAGTGTATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1674  
| | | | |  
QY 1660 ACCAAGATAGTGAAGATTCCTTCAACTCCACCAACAAATGACAGTCTCCATTCACAAAG 1719  
| | | | |  
Db 1675 AGCAAGATAGTGAAGATTCCTTCAACTCCACCAACAAATGACAGTCTCCATTCACAAAG 1734  
| | | | |  
QY 1720 AACCCAAACGCAATCGGAGCCTTAAAGCACTGTAGTGAAGGGCGGCCAGAAAGGATC 1779  
| | | | |  
Db 1735 AACCCAAATGCAATCGAGGCCCAACACCTGTAGTGAAGGGCGGCCAGAAAGGATC 1794  
| | | | |  
QY 1780 CTGACGCGATGAGTGTATCTCTCTCCACGCGGAGAGAACACCTCTGACGAAAGCTG 1839  
| | | | |  
Db 1795 CTGACGCGCTGAGCTTCATCTTCTCCAGGCAAGAGACAGCCTATGATGAAGAGCTG 1854  
| | | | |  
QY 1840 AAGACGCGCTTTCAGAAATGCTACTAGAGCTGGGGGCTTGGAGAGGCTGTAGGT 1899  
| | | | |  
Db 1855 AAGACGCGCTTTCAGAAATGCTACTAGAGCTGGGGGCTTGGAGAGGCTGTAGGT 1914  
| | | | |  
QY 1900 TTCTGCGACCTCTCTGT 1959  
| | | | |  
Db 1915 TTCTGCGACCTCTCTGT 1974  
| | | | |  
QY 1960 GAAGTCAATTTCCCGGTGATTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2019  
| | | | |  
Db 1975 GATGTCAATTTCCCGGTGATTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2034  
| | | | |  
QY 2020 GATGTCAATTTCCCGGTGATTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2079  
| | | | |  
Db 2035 GATGTCAATTTCCCGGTGATTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2094  
| | | | |  
QY 2080 ATGTGTACAGGAGACCATCATCATCAAGCAAGCAATGTGTGAAGGGGGTGGGCAATATC 2139  
| | | | |  
Db 2095 ATGTGTACAGGAGACCATCATCATCAAGCAAGCAATGTGTGAAGGGGGTGGGCAATATC 2154  
| | | | |  
QY 2140 TCAGAAAGTGAAGAGACCGTGAAGACATTTGTGCTGCCGCTCAACAATTCAGTGAACAG 2199  
| | | | |  
Db 2155 TCAGAAAGTGAAGAGACCGTGAAGACATTTGTGCTGCCGCTCAACAATTCAGTGAACAG 2214  
| | | | |  
QY 2200 GTGAACCCAGAGATGCCAAAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2259  
| | | | |  
Db 2215 GTGAACCCAGAGATGCCAAAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2274  
| | | | |  
QY 2260 TCTGAAGAGTGAATGACATTTTGGGAGTACCAACAGGAGATTTGTTTGTGTGAAGACCTCT 2319  
| | | | |  
Db 2275 TCTGAAGAGTGAATGACATTTTGGGAGTACCAACAGGAGATTTGTTTGTGTGAAGACCTCT 2334  
| | | | |  
QY 2320 CTTGAACAGAAATCTTATTTGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2379  
| | | | |  
Db 2335 CTTGAACAGAAATCTTATTTGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2394  
| | | | |  
QY 2380 ACAGGAGATGTGTCAATGATCTTCCAGCTTTGAAAAAGGCAAGATATTGGGGTGTGCATG 2439  
| | | | |

Db 2395 ACAGGAGATGTGTCAATGATCTTCCAGCTTTGAAAAAGGCAAGATATTGGGGTGTGCATG 2454  
| | | | |  
QY 2440 GGGATTTGTGCTGGAGTGTGCCAAGCAAGCTGTGACATGATTTCTTGTGATGACAAAC 2499  
| | | | |  
Db 2455 GGGATTTGTGCTGGAGTGTGCCAAGCAAGCTGTGACATGATTTCTTGTGATGACAAAC 2514  
| | | | |  
QY 2500 TTGGCTGCATCTGTGACCTGTGAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2559  
| | | | |  
Db 2515 TTGGCTGCATCTGTGACCTGTGAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2574  
| | | | |  
QY 2560 TCCATGTCTTACACCTTAAACAAATGAAATTCGGAATTCACCCCTTGTGATATTATTT 2619  
| | | | |  
Db 2575 TCCATGTCTTACACCTTAAACAAATGAAATTCGGAATTCACCCCTTGTGATATTATTT 2634  
| | | | |  
QY 2620 ATTGCAAACTTTCACCTGCGCTGTGGACCTGTGACATTCCTGTGATTTGACTTGGGCACT 2679  
| | | | |  
Db 2635 ATTGCAAACTTTCACCTGCGCTGTGGACCTGTGACATTCCTGTGATTTGACTTGGGCACT 2694  
| | | | |  
QY 2680 GACATGTGTCCCGCATCTCTGTGCTATGAAACAGGCTGAAAGTGAATCAATGAAGAG 2739  
| | | | |  
Db 2695 GACATGTGTCCCGCATCTCTGTGCTATGAAACAGGCTGAAAGTGAATCAATGAAGAG 2754  
| | | | |  
QY 2740 CAGCCCAAAATCCAAAAAGCAAAAATTGTGAACGAGCTGTGATGAGATGGCTTAT 2799  
| | | | |  
Db 2755 CAGCCCAAAATCCAAAAAGCAAAAATTGTGAACGAGCTGTGATGAGATGGCTTAT 2814  
| | | | |  
QY 2800 GGAACGATGTGTATATCCAGGCGCTGTGGAGGCTTCTGACTTATTTGTGATTTGCGCT 2859  
| | | | |  
Db 2815 GGAACGATGTGTATATCCAGGCGCTGTGGAGGCTTCTGACTTATTTGTGATTTGCGCT 2874  
| | | | |  
QY 2860 GAGAACGTTTCTGCGCTTTCACCTGTGTGGCATTCGAGAGACCTGTGATGACCTGTG 2919  
| | | | |  
Db 2875 GAGAACGTTTCTGCGCTTTCACCTGTGTGGCATTCGAGAGACCTGTGATGACCTGTG 2934  
| | | | |  
QY 2920 ATTTAATGATGTGAAGACGCTTACGCGGACGAGTGTGACCTTACAGAGAGAGAAATTTG 2979  
| | | | |  
Db 2935 GTCAACGATGTGAAGACGCTTACGCGGACGAGTGTGACCTTACAGAGAGAGAAATTTG 2994  
| | | | |  
QY 2980 GAGTTCACCTGCGACAGCGCTTCTTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3039  
| | | | |  
Db 2995 GAGTTCACCTGCGACAGCGCTTCTTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3054  
| | | | |  
QY 3040 GTCATGTGAAGACCAAGAAATTTCTGTCTTCACAGAGGAAATGAAGAAATGATTTTA 3099  
| | | | |  
Db 3055 GTCATGTGAAGACCAAGAAATTTCTGTCTTCACAGAGGAAATGAAGAAATGATTTTA 3114  
| | | | |  
QY 3100 ATATTGGCTCTTTGAAGAGACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3159  
| | | | |  
Db 3115 ATATTGGCTCTTTGAAGAGACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3174  
| | | | |  
QY 3160 GGTGACGCTTATGATGATCCCTCAAACTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3219  
| | | | |  
Db 3175 GGTGACGCTTATGATGATCCCTCAAACTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3234  
| | | | |  
QY 3220 TCCCTTCTATCTTGT 3279  
| | | | |  
Db 3235 TCCCTTCTATCTTGT 3294  
| | | | |  
QY 3280 GGTGTGTGTGAAGAAACCTTACTAGACCACTGCGCTGTGACGCGCTGTGAACATTTGT 3339  
| | | | |  
Db 3295 GGTGTGTGTGAAGAAACCTTACTAGACCACTGCGCTGTGACGCGCTGTGAACATTTGT 3354  
| | | | |  
QY 3340 GGCACACACTGACCTTACCCCTTACCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3399  
| | | | |  
Db 3355 GGCACACACTGACCTTACCCCTTACCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3414  
| | | | |  
QY 3400 CTGTACCTGTGTGAAGAAAGCAAAAGCAATGTGGGGATTCAGAGCTGTGAAATGAAGCA 3459  
| | | | |  
Db 3415 CTGTACCTGTGTGAAGAAAGCAAAAGCAATGTGGGGATGTGAGAGCTGTGAAATGAAGCA 3474  
| | | | |  
QY 3460 TGTAGCTGTATGTGGGGCT-----GGGGGAGAGGCTTGGCCGAAAAACCGTGTGACGG 3513  
| | | | |















QY	1201	ATCTCTCTCATTTGATCATGTAGTACGACAAAGTGTGCGGAGAGCTTTGCTGGGCAACCTGTAAAG	1260
Db	1207	ATCTTCTCATTTGATCATGTATCATGTACCCAAACCTGTCCAAAGATGTGTGTGGCCACGTGAACG	1266
QY	1261	GTATGTCTGAGCGCTCACTGGCCAAAGGAGATGGCGAGAAAGAACTGCTCGTGTGAGAACCTTG	1320
Db	1267	GCTGTCTGACACTGACTGATCCCAAGCGCATGGCGAGAGAAAGAACTGCTCGTGTGAGAACTTG	1326
QY	1321	GAAAGTGTGAGACCTTGGGGGTTCACATCCATCTTGTTCGACAAAGACTGGAACCTTG	1380
Db	1327	GAAAGTGTGAGACCTTGGGGGTTCACATCTCTCCATCTGTGCTGACAAAGACCGGAAACCTTG	1386
QY	1381	ACTCAGAAACCGGATGACAGTGGCTCACAATGTGTGTGAGCATCAAAATCATGAAGCTGAC	1440
Db	1387	ACTCAGAAACCGGATGACAGTGGCTCACAATGTGTGTGAGCATCAAAATCATGAAGCTGAC	1446
QY	1441	ACCAAGAGAAATCAGAGTGGGGGTCTCCCTTTGACAGACGTCAGGCCACTGTGGCTCTG	1500
Db	1447	ACCAAGAGAAATCAGAGTGGGGGTCTCTCTTTGACAGACGTCAGGCCACTGTGGCTCTG	1506
QY	1501	TCCAGAAATTCGTGTCTCTGTACACGGGCACTGTTCAGGCTTAACCAAGAAACCTGGCT	1560
Db	1507	TCCAGAAATTCGTGTCTCTGTAAACAGGCAAGTGTTCAGGCTTAACCAAGAAACCTGGCT	1566
QY	1561	ATTCCTTAACGCTGACGTAGCGGAGATGCTTTCGAGTCCGCGCTCTTAAAGTGCATCGAG	1620
Db	1567	ATTCCTTAACGCGGACGTAGCGGAGATGCTCTGAGTGGCGCTCTTAAAGTGCATTTGAA	1626
QY	1621	GTCTGTGTGGCTCCGTGTAGTACATGAGGAGAAAGTATACCAAAATAGTGGAGATTTCCG	1680
Db	1627	GTCTGTGTGGCGCTCCGTGTATGAGATGAGAGGAAATATAGAAATAGTGGAGATTTCCG	1686
QY	1681	TTCAACTCCACAAACAAGTACCAGCTCTCCATTCACAAACACCAAAACCGCATCTGAGTCT	1740
Db	1687	TTCAACTCCACAAACAAGTATCCAGTGTGTCAATTACAAAGAACCCAAATGCATCTGAGCCG	1746
QY	1741	AAGCAGCTGTACTGATGAAAGGGCGGCCCAAGAAAGATCTGTGACCCGATGTATCTATC	1800
Db	1747	AAGCAGCTGTACTGATGAAAGGGCGGCCCAAGAAAGATCTGTGACCGCTGACGCTCATC	1806
QY	1801	CTCTCTCAACGGGAAAGAGACAGCCCTCGAGAAAGAAAGACTGAGAGACGCTTTCAGAAATGCC	1860
Db	1807	CTCTCTCAACGGGAAAGAGACAGCCCTCGAGAAAGAAAGACTGAGAGACGCTTTCAGAAAGCCG	1866
QY	1861	TACCTAGAGCTGGGGGGGCTTTGGAGAGCGTGTCTAGAGTTCTGCCACTCTTTCTGTCT	1920
Db	1867	TACCTAGAGCTGGGGGGGCTTTGGAGAAACGTTGCTAGAGTTCTGCCACTCTCTGTCTCT	1926
QY	1921	GACGACAGTTTCCCGAAGGCTTCCAGTTTGAACCTGATGAAATCAATTTCCCGGTGAT	1980
Db	1927	GACGACAGTTTCCCGAAGGCTTCCAGTTTGAACCTGATGAAATCAATTTTCCCGGTGAT	1986
QY	1981	AAGCTCGCTCGTGGGATCTTATCTCATGATTTGACCCCTCTCGAGCTGTCTGTCCCGAT	2040
Db	1987	AAGCTCGCTCGTGGGATCTTATCTCATGATTTGACCCCTCTCGAGCTGTCTGTCCCGAT	2046
QY	2041	GCTGTGGGCAAAATCCCGAGCGCTGGGATTAAGGTCAATATGTGACAGAGACATATCCA	2100
Db	2047	GCTGTGGGCAAAATCCCGAGCGCTGGGATTAAGGTCAATATGTGACAGAGACATATCCA	2106
QY	2101	ATCAGCGCAAAAGCATTTGTAAGGGGATGGGACCTTATCTCAGAAAGGTAAAGAGACCGTG	2160
Db	2107	ATCAGCGCAAAAGCATTTGTAAGGGGATGGGACCTTATCTCAGAAAGGTAAAGAGACTGTG	2166
QY	2161	GAAAGCATTTGCTGCCCCCTCAACATTTCACTGAAACCAAGTGAACCCCAAGAGATGGCAAG	2220
Db	2167	GAAAGCATTTGCTGCCCCCTCAACATTTCACTGAAACCAAGTGAACCCCAAGAGATGGCAAG	2226
QY	2221	GCGGTGTAGTAATATGGAGAGTCACTTGAAGAAATAGATCTTTGAAGTGTGATGACATTT	2280
Db	2227	GCGGTGTAGTTCACAGGAGTCACTTGAAGAAATAGTACTCTCGAAGAAATCTGATGACATTC	2286















```

/codon_start=2
/product="Similar to ATPase, Na+/K+ transporting, alpha
1a.1 polypeptide"
/db_xref="GI:16307541"
/translation="DDHKLSDLELHRYKTDLSRLTPARAAILARDGPALTPPT
TPKWKRCPLFGGFSMLMLWGLLVCFLVGRATEEEDNDLYGVLSAVIIT
GCEHYRBAKSKKIMESPKMVPPOALVIRGEKMSINAEVWGLDVEVGGRIPA
DLRIISANGCKVNDSSLTGSESPQTRSDPTFNEPLETRNIAFPSTNCVEBTAGIV
YVGRITWGRATLAVSLGEGQTPIAEIEHFHLITGVAVFLVSFLLSLEYTA
LEAVIPLIGIIVANPEGLATVVCILITAKRANRCLVNEAVEITLSTSTIGS
DKTGTILONRMTVAHMFNDQIHEADTENOSGVSFDSATWFLSIAELCKRAVF
OANENPILIKBAVADGASBALKCIEECGSGVEMSEKSKIVEIPNSTNKYOLS
IHNKPNASEPHLLVMKGAPEPILDRCSILHKGEOPLDELEAFQNAVLEFGIGL
ERVIGFCHLLPDEQPEGFOPPTDDNPFYDNLCEFVLIEMIDPPRAVDAVAKCR
SAGIKVIMVTDGHPITAKAIKGVGIISEGETVEDIARINIPNOVNEPDAKCV
HGSOLDKMTSELEDDILRYHTEIVFARTSPQCKLIYEGCQROGAIVAVTDGVNDP
ALKKADIGVAMGIVSGDVSQKQADMLILDDNFASIVTGEGRGLFDMLKKSIVYTL
SNIPETPELFIITANIPLEPLGTVILICIDGTMBPISLAYEASEDIDKROPNP
KTDKLVNRPILSMAYGOIGMIOALGGFTYVILIAENGLPPLLGISETMDRWDND
VEDSYGOQWITYEOKIYEFCTHTAFVSVIVVQNAIDVICTRNSVFOQMKKILLI
FGLFEETLAAFLSYCPQMGALRMYPLKPTMFCAPFYSLLIFVDEVRLLIIRRP
GGWKEKEYYY"

```

BASE COUNT 795 a 859 c 881 g 762 t

ORIGIN

Query Match 82.9% Score 3014.6; DB 10; Length 3297;  
 Best Local Similarity 95.1%; Pred. No. 0; Mismatches 154; Indels 6; Gaps 1;  
 Matches 3123; Conservative 0;

```

360 GGAGACCATTAACACAGCCTGGATGACATCCATGTAATAACGAACAGATTGAGCG 419
Db 1 GGACACCATTAACACAGCCTGGATGACATCCATGTAATAACGAACAGATTGAGCG 60
QY 420 AGGCTTAACACCGGAGGCGCGCTGAGATCCTGGCTGGGATGGCCCAAGCCCTTAC 479
Db 61 AGGATTAACACCGTGAAGGCGCGCTGAGATCCTGGCGGAGTGGCCCAAGCCCTTAC 120
QY 480 GCGGCTGCACTACTCGGAGTGGTAAATTTCTGGGACGCTTTGGTGGCTTCTC 539
Db 121 ACCCCTCCCACTACTCCCGGATGGTGAAATTTCTGCCGCACTCTTTGAGAGCTTTC 180
QY 540 CATGTTACTGTGGATTGAGCAATCTTTGTTCTTGCTTATGGCATCCGAAGTCTAC 599
Db 181 CATGTTACTGTGGATTGAGGAGCATTTCTTGTCTTGCTTATGGCATCCGAAGTCTAC 240
QY 600 AGAAGAGGAACCAACAAATGATGATGTAACCTCGGAGTGGTGTCTGTCTGCTGCTAC 659
Db 241 AGAAGAGGAACCGCAACAAAGATGATCTGATCTCGGAGTGGTGTCTCTGCTGCTAC 300
QY 660 CATTAAGCTGCTTTCTCTATTTATCAGAGCAAAAAGCTCCAGATCATGGAATCTTT 719
Db 301 CATTAAGCTGCTTTCTCTATTTATCAGAAACCAAAAAGCTCCAGATCATGGAATCTTT 360
QY 720 CAAGAAATGTCCTCTCAGCAAGCCCTCGTATTCGAAATGAGAGAAATGACATCAA 779
Db 361 CAAGAAATGTCCTCTCAGCAAGCCCTCGTATTCGAAATGAGAGAAATGACATCAA 420
QY 780 CCGAGAGATGTCGTCGTTGATGATCTGGTGAAGTGAAGGCGGAGACCGAATCCCTGC 839
Db 421 TGGCGAGATGTCGTCGTTGATGATCTGGTGAAGTGAAGGCGGAGACCGAATCCCTGC 480
QY 840 TGATCTCAGATCATATCTGCAAAATGCTGCAAGTGGATTAATCTCTCACTGCTGTA 899
Db 481 TGATCTCAGATCATATCTGCAAAATGCTGCAAGTGGATTAATCTCTCACTGCTGTA 540
QY 900 ATAGAAACCCGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 959
Db 541 ATAGAAACCCGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 600
QY 960 CATTCGCTTCTTCAACCAATGTTGTAAGAACTGCAAGTGGCTGTTGTAACAC 1019
Db 601 CATTCGCTTCTTCAACCAATGTTGTAAGAACTGCAAGTGGCTGTTGTAACAC 660

```

```

QY 1020 TGGGATGCAACCGGATGGGAGGATGCGCACCTTGCTTCTGAGGCTGGAAGCGGCTT 1079
Db 661 TGGGATGCAACCGGATGGGAGGATGCGCACCTTGCTTCTGAGGCTGGAAGCGGCTT 720
QY 1080 GACCCCAATGCTGAAGAAATGAGCACTTCATGCACTCATACGGGTGTGGCGCTGTT 1139
Db 721 GACCCCAATGCTGAAGAAATGAGCACTTCATGCACTCATACGGGTGTGGCGCTGTT 780
QY 1140 CCTGGGGGTGTCTTTCTTCTTCTCTCTGATTCCTTGATACCTGGCTGAGGCTGT 1199
Db 781 CCTGGGGGTGTCTTTCTTCTTCTCTGATTCCTTGATACCTGGCTGAGGCTGT 840
QY 1200 CATTCCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259
Db 841 CATTCCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 1260 GGTATGTCTGACGCTCACTGCAAGCGCATGGCGAGAAAGAACTGCTGGTGAAGACT 1319
Db 901 GGTATGTCTGACGCTCACTGCAAGCGCATGGCGAGAAAGAACTGCTGGTGAAGACT 960
QY 1320 GGAAGCTGTGAGACCTTGGGGTCCACATCCACATCTGCTCGACAAAGCTGGAACCT 1379
Db 961 GGAAGCTGTGAGACCTTGGGGTCCACATCTGCTCGACAAAGCTGGAACCTT 1020
QY 1380 GACTCAGAACCGGATGACAGTGGCTCACAATGCTTGAACAATCAATCCATGAGCTGA 1439
Db 1021 GACTCAGAACCGGATGACAGTGGCTCACAATGCTTGAACAATCCATGAGCTGA 1080
QY 1440 CACACAGAAATCAGAGTGGGTCTCTTTGACAAAGAGTCAAGCACTGCTGCTCT 1499
Db 1081 CACACAGAAATCAGAGTGGGTCTCTTTGACAAAGAGTCAAGCACTGCTGCTCT 1140
QY 1500 GTCCAGAAATGCTGCTCTGTAAACAGGCGAGTGTTCAGGCTAACCAAGAAACCTGCC 1559
Db 1141 GTCCAGAAATGCTGCTCTGTAAACAGGCGAGTGTTCAGGCTAACCAAGAAACCTGCC 1200
QY 1560 TATCTTAAAGCTGACGTAAGCGGAGATGCTTCGAGTGGGCGCTCTTAAAGTGCATGA 1619
Db 1201 TATCTTAAAGCGGAGTGAAGCGGAGATGCTTCGAGTGGGCGCTCTTAAAGTGCATGA 1260
QY 1620 GGTCTGCTGGCTCGCTGATGATGAGAGTGAAGGAGTACACCAAGATGAGAGATTCC 1679
Db 1261 AGTCTGCTGGCTCGCTGATGATGAGAGTGAAGGAGTACACCAAGATGAGAGATTCC 1320
QY 1680 CTTCAACTCCACCAACAGTACCAAGTCTCCATTCACAAAGCCCAACCGCATGAGGCC 1739
Db 1321 CTTCAACTCCACCAACAGTACCAAGTCTCCATTCACAAAGCCCAACCGCATGAGGCC 1380
QY 1740 TAAAGCACTGCTGATGATGAAAGGCGCCCAAGAAAGATCTCTGGACCGATGCTTAT 1799
Db 1381 CAAAGCACTGCTGATGATGAAAGGCGCCCAAGAAAGATCTCTGGACCGATGCTTAT 1440
QY 1800 CTTCTCTCAACGAGAGAGAGCGCTTGAAGCAAGAGCTGAAGAGCGCTTTGAGAAATGC 1859
Db 1441 CTTCTCTCAACGAGAGAGAGCGCTTGAAGCAAGAGCTGAAGAGCGCTTTGAGAAATGC 1500
QY 1860 CTACCTAGAGCTGGGCGCTTGAAGAGGCTGTAAGCTGTAAGCTGTAAGCTGTAAGCT 1919
Db 1501 CTACCTAGAGCTGGGCGCTTGAAGAGGCTTGAAGCTGTAAGCTGTAAGCTGTAAGCT 1560
QY 1920 TGAAGCAAGATTTCCGGAAGGCTTCAAGTTGACCTGATGAAGTCAATTTCCCGTGA 1979
Db 1561 TGAAGCAAGATTTCCGGAAGGCTTCAAGTTGACCTGATGAAGTCAATTTCCCGTGA 1620
QY 1980 TAACTGCTGCTGCTGCTTATCTTCATGATGATGATGATGATGATGATGATGATG 2039
Db 1621 TAACTGCTGCTGCTGCTTATCTTCATGATGATGATGATGATGATGATGATGATG 1680
QY 2040 TGTCTGAGCAATTCGCGAGCGCTGGGATTAAGTATGATGATGATGATGATGATGAT 2099
Db 1681 TGTCTGAGCAATTCGCGAGCGCTGGGATTAAGTATGATGATGATGATGATGATGATG 1740

```















Disclosure; Page 59-63; 72pp; English.











FT mat\_peptide 241..3636  
 FT /tag= C  
 XX /label= Mature\_(Na,K)-ATPase\_799/801NKA  
 PN US6309874-B1.  
 XX 30-OCT-2001.  
 PD 04-JUN-1998; 98US-0090535.  
 XX 04-JUN-1997; 97US-048601P.  
 PR (KARO-) KAROLINSKA INNOVATIONS AB.  
 PA Belusa R;  
 PI WPI, 2002 040238/05.  
 DR P-PDR: AAT10501.  
 XX New nucleic acid encoding an ouabain-resistant (Na,K)-ATPase for use as  
 PT a selectable marker for transformed mammalian cells, particularly  
 PT suitable when selecting cells for transplant into a patient  
 XX Disclosure: Column 48-50; 30pp; English.  
 XX  
 CC The invention relates to a nucleic acid encoding an ouabain-resistant  
 CC (Na,K)-ATPase termed 799/801NKA, a vector comprising the nucleic  
 CC acid and a cell comprising the vector. The nucleic acid is useful as a  
 CC selective marker to screen for transformed cells in gene therapy. Unlike  
 CC prior art selective markers, the gene is already expressed in all  
 CC mammalian cells and the minor mutation is more natural than adding an  
 CC artificial gene or making major mutations so the invention is more  
 CC suitable for selecting cells for transplantation into a patient.  
 CC The present sequence encodes wild-type rat (Na,K)-ATPase protein  
 CC used as the basis for the mutated proteins of the invention.  
 XX  
 SO Sequence 3636 BP; 850 A; 968 C; 995 G; 823 T; 0 other;  
 Query Match 99.9%; Score 3611.2; DB 24; Length 3636;  
 Best Local Similarity 99.9%; Pval No 0;  
 Matches 3633; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 481 CCCCTCCACTACTCCCGAGTGGGCAATTCGTGCGGACCTGTGGGCTTCTCC 540  
 DB 481 CCCCTCCACTACTCCCGAGTGGGCAATTCGTGCGGACCTGTGGGCTTCTCC 540  
 QY 541 ATGTTACTGTGATGAGACCATTTCTTGTCTTGTGCTTATGAGATGCTGACA 600  
 DB 541 ATGTTACTGTGATGAGACCATTTCTTGTCTTGTGCTTATGAGATGCTGACA 600  
 QY 601 GAAAGAGAAACCAAAATGATGATCTGTACCTCGGAGTGTGTCTGTCTGTCT 660  
 DB 601 GAAAGAGAAACCAAAATGATGATCTGTACCTCGGAGTGTGTCTGTCTGTCT 660  
 QY 661 ATTAAGCTGCTGTTCTCCATATATCAAGAGAAAGAAAGCTCCAGATCTGGAATCTTCG 720  
 DB 661 ATTAAGCTGCTGTTCTCCATATATCAAGAGAAAGAAAGCTCCAGATCTGGAATCTTCG 720  
 QY 721 AAGAAATGATGCTGCTGAGCAAGCCCTGTGATTTGAAATGAGAGAAATGAGCATCAAC 780  
 DB 721 AAGAAATGATGCTGCTGAGCAAGCCCTGTGATTTGAAATGAGAGAAATGAGCATCAAC 780  
 QY 781 GCAAGAGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 DB 781 GCAAGAGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 QY 841 GATCTCAGATCATATCTGCAATGAGTGCAGAGTGAATCTCCCTCACTCACTGATGAA 900  
 DB 841 GATCTCAGATCATATCTGCAATGAGTGCAGAGTGAATCTCCCTCACTCACTGATGAA 900  
 QY 901 TGAAACCCCAAGACTGCTGCTCCCGATTTTCAACAAAGAGACCCCTTGAGACAAAGAAC 960  
 DB 901 TGAAACCCCAAGACTGCTGCTCCCGATTTTCAACAAAGAGACCCCTTGAGACAAAGAAC 960  
 QY 961 ATTGCTGCTGCTCAACCAAGTGTGTAAGAACTGACAGTGGATGCTGTGTGATCACT 1020  
 DB 961 ATTGCTGCTGCTCAACCAAGTGTGTAAGAACTGACAGTGGATGCTGTGTGATCACT 1020  
 QY 1021 GGGGATGCAAGCTGATGAGGAGATGAGCAAGTGCACCTTGTCTTGTGGCTGAGAGCGGCTTG 1080  
 DB 1021 GGGGATGCAAGCTGATGAGGAGATGAGCAAGTGCACCTTGTCTTGTGGCTGAGAGCGGCTTG 1080  
 QY 1081 ACCCCATGCTGAAAGAAATGAGACCTTCATCCATCATACAGGATGAGGCTGCTGCTG 1140  
 DB 1081 ACCCCATGCTGAAAGAAATGAGACCTTCATCCATCATACAGGATGAGGCTGCTGCTG 1140  
 QY 1141 CTGGGGGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
 DB 1141 CTGGGGGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
 QY 1201 ATCTTCTGATTTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
 DB 1201 ATCTTCTGATTTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
 QY 1261 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 DB 1261 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 QY 1321 GAAAGCTGTGAGACCTTGGGGTCCATGACATGACATGACATGACATGACATGACATGAC 1380  
 DB 1321 GAAAGCTGTGAGACCTTGGGGTCCATGACATGACATGACATGACATGACATGACATGAC 1380  
 QY 1381 ACTCAGAACCGGATGACAGTGTGCTCATGTGCTTGTGACATGATGATGATGATGATGAT 1440  
 DB 1381 ACTCAGAACCGGATGACAGTGTGCTCATGTGCTTGTGACATGATGATGATGATGATGAT 1440  
 QY 1441 ACCACAGAAATCAAGAGTGGGGTCTCTTTGACAGAGCTGACGACCTGTTGCTGCTG 1500  
 DB 1441 ACCACAGAAATCAAGAGTGGGGTCTCTTTGACAGAGCTGACGACCTGTTGCTGCTG 1500  
 QY 1501 TCCAGATTTGCTGCTCTGTAAACAGAGGAGTGTTCAGGCTTAAACCAAGAAACCTGCT 1560  
 DB 1501 TCCAGATTTGCTGCTCTGTAAACAGAGGAGTGTTCAGGCTTAAACCAAGAAACCTGCT 1560



b	2641	CTGGGACCGTACCATCTCTGCATTGACTTTGGACACTGATCGTATCCGGCATCTCT	2700
y	2701	CTGGCCCTATGAACAGGCTGAAAGTGACATCATGAAGGCGACCCAGAAATCCCAAAACG	2760
b	2701	CTGGCCCTATGAACAGGCTGAAAGTGACATCATGAAGGCGACCCAGAAATCCCAAAACG	2760
y	2761	GACAAACTTGTGAACGAGCGCTGTGATTCAGCATGGCTATGGAACGATCGGATGATCCAG	2820
b	2761	GACAAACTTGTGAACGAGCGCTGTGATTCAGCATGGCTATGGAACGATCGGATGATCCAG	2820
y	2821	GGCCTGGAGAGGCTCTTCACTATTATTGTGATTTCTGCGTGAACCGGTTTCTGCCCTTT	2880
b	2821	GGCCTGGAGAGGCTCTTCACTATTATTGTGATTTCTGCGTGAACCGGTTTCTGCCCTTT	2880
y	2881	CACCTGTGGGCATCCGAAGACCTTGGCATGACCCCTGGATTAAAGATGGAGACAGC	2940
b	2881	CACCTGTGGGCATCCGAAGACCTTGGCATGACCCCTGGATTAAAGATGGAGACAGC	2940
y	2941	TAGGGGACAGAGTGACCTACGACGACGAGAAAGATTGTGAGTTCACTGCCACAGGCC	3000
b	2941	TAGGGGACAGAGTGACCTACGACGACGAGAAAGATTGTGAGTTCACTGCCACAGGCC	3000
y	3001	TTCTTTTCAGTATCGTGTGATGTGACGCTGGCTGACCTGGTATCGCAAGCCAGAAAG	3060
b	3001	TTCTTTTCAGTATCGTGTGATGTGACGCTGGCTGACCTGGTATCGCAAGCCAGAAAG	3060
y	3061	AATTTCGTCTTCAGCGAGGGAATGAAGAACAAAGTCTTAATATTGGCCTCTTTGAAG	3120
b	3061	AATTTCGTCTTCAGCGAGGGAATGAAGAACAAAGTCTTAATATTGGCCTCTTTGAAG	3120
y	3121	ACAGCTCTTGCTGCTTCTGTCTACTGCTCCCTGGGATGGGTGACACCCTTAGAGATGT	3180
b	3121	ACAGCTCTTGCTGCTTCTGTCTACTGCTCCCTGGGATGGGTGACACCCTTAGAGATGT	3180
y	3181	CCCCCAAACTCACTAGTGGTGTGTGTGACCTTCCCACTACCTCTTCTCATCTTGTGTAT	3240
b	3181	CCCCCAAACTCACTAGTGGTGTGTGTGACCTTCCCACTACCTCTTCTCATCTTGTGTAT	3240
y	3241	GACGAGTGTGGAGCTCATCATGAGCGACGCCCTGGCGGCTGGGTGAGAAAGAAACC	3300
b	3241	GACGAGTGTGGAGCTCATCATGAGCGACGCCCTGGCGGCTGGGTGAGAAAGAAACC	3300
y	3301	TACTACTAGCCCACTGCGCTGACGCGGTGGGAACTTGTGTGTGTGTGTGTGTGTGTGT	3360
b	3301	TACTACTAGCCCACTGCGCTGACGCGGTGGGAACTTGTGTGTGTGTGTGTGTGTGTGT	3360
y	3361	TACCCCCCTTTGT	3420
b	3361	TACCCCCCTTTGT	3420
y	3421	CCAAAGCATGTGGGATCCGACGCTCTGGAATGAAGATGTAGCTGTAAATGGGGGCGG	3480
b	3421	CCAAAGCATGTGGGATCCGACGCTCTGGAATGAAGATGTAGCTGTAAATGGGGGCGG	3480
y	3481	GGGAGAGGCTGCCGGAAGAACCGGTGGACGGGACGACACGGGGGAAAGGTTTATGTG	3540
b	3481	GGGAGAGGCTGCCGGAAGAACCGGTGGACGGGACGACACGGGGGAAAGGTTTATGTG	3540
y	3541	CGTTTGT	3600
b	3541	CGTTTGT	3600
y	3601	GGATTTTTCAAATTAAGATGGCTATTATTAAGGAA	3636
b	3601	GGATTTTTCAAATTAAGATGGCTATTATTAAGGAA	3636

RESULT 3  
AAV84294  
ID AAV84294 standard; DNA; 3636 BP  
XX  
AC AAV84294;







Db 1381 ACTCAGAACCGGATGACAGTGGCTCACATGTGGTTTGACAAATCAATTCATGAAGCTGAC 1440  
Qy 1441 ACCACAGAGAAATCAGAGTGGGGTCTCCTTTGACAAAGCGTCAGCCACTGGTGTGCTG 1500  
Db 1441 ACCACAGAGAAATCAGAGTGGGGTCTCCTTTGACAAAGCGTCAGCCACTGGTGTGCTG 1500  
Qy 1501 TCCAGAAATTCCTGCTCTCTGTATACAGGGCAGTGTTCAGAGCTAAACAAAGAAACCTGCTT 1560  
Db 1501 TCCAGAAATTCCTGCTCTCTGTATACAGGGCAGTGTTCAGAGCTAAACAAAGAAACCTGCTT 1560  
Qy 1561 ATCTTAAAGGTGAGTAGAGCGGGAGATGCTTCGAGTGGGGCTCTTAAAGTGCATCGAG 1620  
Db 1561 ATCTTAAAGGTGAGTAGAGCGGGAGATGCTTCGAGTGGGGCTCTTAAAGTGCATCGAG 1620  
Qy 1621 GTCTGCTGTGGCTCCGTGATGAGATGAGGAGAAATAACCAAGATAGTGGAGATTCCC 1680  
Db 1621 GTCTGCTGTGGCTCCGTGATGAGATGAGGAGAAATAACCAAGATAGTGGAGATTCCC 1680  
Qy 1681 TTCAATCTCCACCAACAGTACAGTCTTCATTTCAAGAAACCCAAAGCATCGAGAGCT 1740  
Db 1681 TTCAATCTCCACCAACAGTACAGTCTTCATTTCAAGAAACCCAAAGCATCGAGAGCT 1740  
Qy 1741 AAGGACCTGGTAGTATGAAGGGGCGCCGAGAAAGATCTGAGACCGATGCAATTCTATC 1800  
Db 1741 AAGGACCTGGTAGTATGAAGGGGCGCCGAGAAAGATCTGAGACCGATGCAATTCTATC 1800  
Qy 1801 CTCTCTCCACGGCAGAGCAGCCCTGAGCGAAGAGCTGAAAGACGCTTTCAAGATGCC 1860  
Db 1801 CTCTCTCCACGGCAGAGCAGCCCTGAGCGAAGAGCTGAAAGACGCTTTCAAGATGCC 1860  
Qy 1861 TACCTAGAGCTGGGGGGGCTTGGAGAGGGTGTGCTAGAGTTCTGGCACCCTCTTCTGCT 1920  
Db 1861 TACCTAGAGCTGGGGGGGCTTGGAGAGGGTGTGCTAGAGTTCTGGCACCCTCTTCTGCT 1920  
Qy 1921 GACGAAACAGTTTCCGAAAGGCTTCCAGTTTGAACACTGATGAAGTCAATTTCCCGGTGAT 1980  
Db 1921 GACGAAACAGTTTCCGAAAGGCTTCCAGTTTGAACACTGATGAAGTCAATTTCCCGGTGAT 1980  
Qy 1981 AACCTCTGCTTGGTGGGTCTTATCTCCATGATTTGACCCCTCGAGCTGTGCCCGAT 2040  
Db 1981 AACCTCTGCTTGGTGGGTCTTATCTCCATGATTTGACCCCTCGAGCTGTGCCCGAT 2040  
Qy 2041 GCTGTGGGCAAAATGCGTAAGGCTGGGATTAAGTCAATGCTCAACAGAGCAATCCA 2100  
Db 2041 GCTGTGGGCAAAATGCGTAAGGCTGGGATTAAGTCAATGCTCAACAGAGCAATCCA 2100  
Qy 2101 ATCAACAGCCAAAGCCATTGCTAAAGGGGGTGGGATTAATCTCAGAGGTAAACGAGCCGTG 2160  
Db 2101 ATCAACAGCCAAAGCCATTGCTAAAGGGGGTGGGATTAATCTCAGAGGTAAACGAGCCGTG 2160  
Qy 2161 GAAGCATTGCTGCGCGGCTCAACATTTCCAGTGAACCGAGTGAACCCAGAGATGCCAAG 2220  
Db 2161 GAAGCATTGCTGCGCGGCTCAACATTTCCAGTGAACCGAGTGAACCCAGAGATGCCAAG 2220  
Qy 2221 GCTGTGTAGTATCATGTGACGTGAAGGACATGACCTCTGAGAGCTGTGATGAAT 2280  
Db 2221 GCTGTGTAGTATCATGTGACGTGAAGGACATGACCTCTGAGAGCTGTGATGAAT 2280  
Qy 2281 TTGGGGTACACAGCGAGATTTGCTTTGCTAGGACCTCTCTCAAGAGATCATCAT 2340  
Db 2281 TTGGGGTACACAGCGAGATTTGCTTTGCTAGGACCTCTCTCAAGAGATCATCAT 2340  
Qy 2341 GTGAGAGGCTGCGAGCGGAGGCTGCTATGCTGCTTACAGGGATGCTGTCAATGAC 2400  
Db 2341 GTGAGAGGCTGCGAGCGGAGGCTGCTATGCTGCTTACAGGGATGCTGTCAATGAC 2400  
Qy 2401 TCTCCAGCTTTGAAAAAGGAGATATTGGGGTTGCTGATGGGATTTGTTGGTCCGAGTGTG 2460  
Db 2401 TCTCCAGCTTTGAAAAAGGAGATATTGGGGTTGCTGATGGGATTTGTTGGTCCGAGTGTG 2460  
Qy 2461 TCCAAAGCAAGTGTGATGATTTTCTGTGATGAACCTTTGCTCATGTGATGGA 2520  
Db 2461 TCCAAAGCAAGTGTGATGATTTTCTGTGATGAACCTTTGCTCATGTGATGGA 2520

Qy 2521 GTAGAGAAAGTGTCTGATATTGTTGATTAACCTGAGAAATCCATTTGCTTACACCTACA 2580  
Db 2521 GTAGAGAAAGTGTCTGATATTGTTGATTAACCTGAGAAATCCATTTGCTTACACCTACA 2580  
Qy 2581 AGTAAATTCGGAATATCACCCCTTCTTGATATTATTTATTTGCAACATTCACCTGCC 2640  
Db 2581 AGTAAATTCGGAATATCACCCCTTCTTGATATTATTTATTTGCAACATTCACCTGCC 2640  
Qy 2641 CTGGGACCGTGACCATCTCTGATTAACCTTGGGACATGACATGATGCTCCGACATCT 2700  
Db 2641 CTGGGACCGTGACCATCTCTGATTAACCTTGGGACATGACATGATGCTCCGACATCT 2700  
Qy 2701 CTGGGCTATGAACAGCTGAAAGTGACATCATGAAGAGCAGCCAGAAATCCAAAACG 2760  
Db 2701 CTGGGCTATGAACAGCTGAAAGTGACATCATGAAGAGCAGCCAGAAATCCAAAACG 2760  
Qy 2761 GACAAACTTGTGAACAGCCTGTGATCAGCATGGCTATGACACATGGTATGATCCAG 2820  
Db 2761 GACAAACTTGTGAACAGCCTGTGATCAGCATGGCTATGACACATGGTATGATCCAG 2820  
Qy 2821 GCCCTGGAGGCTCTTCACTATTATTTGATTTCTGGCTGAGAAAGGTTTCTGCTGCTTT 2880  
Db 2821 GCCCTGGAGGCTCTTCACTATTATTTGATTTCTGGCTGAGAAAGGTTTCTGCTGCTTT 2880  
Qy 2881 CACTGTTGAGCATCCGAGAGACCTGGAGTGAACCGCTGGATTATGATGTGAGAGACG 2940  
Db 2881 CACTGTTGAGCATCCGAGAGACCTGGAGTGAACCGCTGGATTATGATGTGAGAGACG 2940  
Qy 2941 TAGGGCAGCATGAGACCTACAGAGAGAGAAATTTGAGAGTTTACCTGGCAACAGGCG 3000  
Db 2941 TAGGGCAGCATGAGACCTACAGAGAGAGAAATTTGAGAGTTTACCTGGCAACAGGCG 3000  
Qy 3001 TTCTTTGAGTATGAGTGTGATGAGAGTGGGCTGACTGGTCACTGCAAGACCAAGAG 3060  
Db 3001 TTCTTTGAGTATGAGTGTGATGAGAGTGGGCTGACTGGTCACTGCAAGACCAAGAG 3060  
Qy 3061 AATTCTGTCTTCAGAGGAAATGAAGAACAAAGTCTTAATTTGAGGCTCTTTGAAGAG 3120  
Db 3061 AATTCTGTCTTCAGAGGAAATGAAGAACAAAGTCTTAATTTGAGGCTCTTTGAAGAG 3120  
Qy 3121 ACAGCTCTGAGCTTCTCTGCTCACTGAGCCCTGGAGTGGGTCAGCCCTTAGAGATGAT 3180  
Db 3121 ACAGCTCTGAGCTTCTCTGCTCACTGAGCCCTGGAGTGGGTCAGCCCTTAGAGATGAT 3180  
Qy 3181 CCGCTCAACCTTATGTTGTGCTTCTGCTTCCCTCACTCCCTTCTCATCTGCTGTAT 3240  
Db 3181 CCGCTCAACCTTATGTTGTGCTTCTGCTTCCCTCACTCCCTTCTCATCTGCTGTAT 3240  
Qy 3241 GAGGAGGTGCGAAGCTCATCATCAGCGAGCGCCCTGGGGCTGGGAGAAAGAAC 3300  
Db 3241 GAGGAGGTGCGAAGCTCATCATCAGCGAGCGCCCTGGGGCTGGGAGAAAGAAC 3300  
Qy 3301 TACTACTAGCCACTGCGCTGCAAGCTCTGAGGCTGGAACCTCTGCTGAGGAAAGCA 3360  
Db 3301 TACTACTAGCCACTGCGCTGCAAGCTCTGAGGCTGGAACCTCTGCTGAGGAAAGCA 3360  
Qy 3361 TACCCCTCTTTGTGATCTTATGATCTTTGAGCTTGGAGCTTGAACCTCTGATGAGAAAGCA 3420  
Db 3361 TACCCCTCTTTGTGATCTTATGATCTTTGAGCTTGGAGCTTGAACCTCTGATGAGAAAGCA 3420  
Qy 3421 CCAAGCATGTGGGATTCAGACGCTCTGGAATGAAGCATGTAGCTGTATGGGGGGGG 3480  
Db 3421 CCAAGCATGTGGGATTCAGACGCTCTGGAATGAAGCATGTAGCTGTATGGGGGGGG 3480  
Qy 3481 GGGAGAGGCTGCCGAAAAACACCGTGAAGGGGAGACAGCGGGGAAAGTTATATG 3540  
Db 3481 GGGAGAGGCTGCCGAAAAACACCGTGAAGGGGAGACAGCGGGGAAAGTTATATG 3540  
Qy 3541 CTTTGTGTTTGTAAAAAGAAAAACCTGGAAAGACTGAAAGATTTGTTTATATCT 3600  
Db 3541 CTTTGTGTTTGTAAAAAGAAAAACCTGGAAAGACTGAAAGATTTGTTTATATCT 3600







```

1261 GATATCTGTAACGCTCACTGTCAGAGGCGATGCGAGAGAACTGCTGTGAGAACCTG 1320
1261 GATATCTGTAACGCTCACTGTCAGAGGCGATGCGAGAGAACTGCTGTGAGAACCTG 1320
1321 GAAGCTGTGAGAACCTGTGGGCTCCACATCCACCATCTGCTCCGCAAGACCTGGAACCTG 1380
1321 GAAGCTGTGAGAACCTGTGGGCTCCACATCCACCATCTGCTCCGCAAGACCTGGAACCTG 1380
1381 ACTCAGAACCCGATGAGAGTGGGCTCCTCCTTTGACAAAGAGTGGGCTGCTGCTG 1500
1381 ACTCAGAACCCGATGAGAGTGGGCTCCTCCTTTGACAAAGAGTGGGCTGCTGCTG 1500
1441 ACCAGAGAGATCAGAGTGGGCTCCTCCTTTGACAAAGAGTGGGCTGCTGCTG 1500
1441 ACCAGAGAGATCAGAGTGGGCTCCTCCTTTGACAAAGAGTGGGCTGCTGCTG 1500
1501 TCCAGAGATGCTGCTCCTGTAACAGGCGACATGTTGAGCTAACCAAGAAACCTGCT 1560
1501 TCCAGAGATGCTGCTCCTGTAACAGGCGACATGTTGAGCTAACCAAGAAACCTGCT 1560
1561 ATCCTTAAAGCGTGGAGTAGCGGAGATGCTTCCAGATGCTGGGCTCTTAAAGTGCATCGAG 1620
1561 ATCCTTAAAGCGTGGAGTAGCGGAGATGCTTCCAGATGCTGGGCTCTTAAAGTGCATCGAG 1620
1621 GTCTGCTGTGGCTCCGTGATGAGATGAGGAGAGATCACCAGATAGTGGAGATTC 1680
1621 GTCTGCTGTGGCTCCGTGATGAGATGAGGAGAGATCACCAGATAGTGGAGATTC 1680
1681 TTCAACTCCACCAACCAATGATCAGCTCTCATTCAAGAAACCCAAAGCAGATCGAGCT 1740
1681 TTCAACTCCACCAACCAATGATCAGCTCTCATTCAAGAAACCCAAAGCAGATCGAGCT 1740
1741 AAGCACTGCTAGTATGATAAGGCGGCGCCAGAAAGATCCTGAGCCGATGAGCTCTATC 1800
1741 AAGCACTGCTAGTATGATAAGGCGGCGCCAGAAAGATCCTGAGCCGATGAGCTCTATC 1800
1801 CTCCTCCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
1801 CTCCTCCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
1861 TAACCTAGAGCTGGGAGGCTTGAAGAGCTGTGCTAGGTTCTGCGCAGCTCTCTGCT 1920
1861 TAACCTAGAGCTGGGAGGCTTGAAGAGCTGTGCTAGGTTCTGCGCAGCTCTCTGCT 1920
1921 GACGAACTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
1921 GACGAACTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
1981 AACCTGCTGCTGGGAGTCTATCTTCATGATGAGAGCTGCTGAGCTGCTGCTGCTGCTGAT 2040
1981 AACCTGCTGCTGGGAGTCTATCTTCATGATGAGAGCTGCTGAGCTGCTGCTGCTGCTGAT 2040
2041 GCTGTGGGCAATTCGCGAGCGCTGGGATTAAGGTCATGCTGACAGAGAGAGAGAGAGAG 2100
2041 GCTGTGGGCAATTCGCGAGCGCTGGGATTAAGGTCATGCTGACAGAGAGAGAGAGAGAG 2100
2101 ATCAGAGCCAAAGCATTGCTAAAGGAGTGGGAGTATCTCAGAGAGTAAAGAGAGAGAGAG 2160
2101 ATCAGAGCCAAAGCATTGCTAAAGGAGTGGGAGTATCTCAGAGAGTAAAGAGAGAGAGAG 2160
2161 GAAAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
2161 GAAAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
2221 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
2221 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
2281 TTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
2281 TTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
2341 GTGAGAGAGCTGCCAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
2341 GTGAGAGAGCTGCCAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400

```







Db 901 TCAGAACCCGAGCTCGGTCCCGGATTTCAAAACGAGAACCCCTTGAGACAGGAGAC 960  
QY ATTGCTTTCTTCTCAACCACTGTGTGAAGAACTGCAGTGCATCGTGTGTAACACT 1020  
Db 961 ATTGCTTTCTTCTCAACCACTGTGTGAAGAACTGCAGTGCATCGTGTGTAACACT 1020  
QY ATTGCTTTCTTCTCAACCACTGTGTGAAGAACTGCAGTGCATCGTGTGTAACACT 1020  
Db 1021 GAGGATGACCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
QY GAGGATGACCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
Db 1021 GAGGATGACCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
QY GAGGATGACCGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
Db 1081 ACCGCTATGCTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAAT 1140  
QY ACCGCTATGCTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAAT 1140  
Db 1081 ACCGCTATGCTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAAT 1140  
QY ACCGCTATGCTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAAT 1140  
Db 1141 CTGGGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
QY CTGGGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
Db 1201 ATCTTCTCATTTGATATCATGTAGCCAACTGCGGAAAGGTTTGCTGCGACCGTCA 1260  
QY ATCTTCTCATTTGATATCATGTAGCCAACTGCGGAAAGGTTTGCTGCGACCGTCA 1260  
Db 1201 ATCTTCTCATTTGATATCATGTAGCCAACTGCGGAAAGGTTTGCTGCGACCGTCA 1260  
QY ATCTTCTCATTTGATATCATGTAGCCAACTGCGGAAAGGTTTGCTGCGACCGTCA 1260  
Db 1261 GTATGCTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAATCA 1320  
QY GTATGCTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAATCA 1320  
Db 1321 GAACTGTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAATCA 1380  
QY GAACTGTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAATCA 1380  
Db 1321 GAACTGTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAATCA 1380  
QY GAACTGTGTAAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAATCA 1380  
Db 1381 ACTGAGAAACCGGATGACAGTGGCTTCAATGATGATGATGATGATGATGATGATG 1440  
QY ACTGAGAAACCGGATGACAGTGGCTTCAATGATGATGATGATGATGATGATGATG 1440  
Db 1381 ACTGAGAAACCGGATGACAGTGGCTTCAATGATGATGATGATGATGATGATGATG 1440  
QY ACTGAGAAACCGGATGACAGTGGCTTCAATGATGATGATGATGATGATGATGATG 1440  
Db 1441 ACCAGAGAAATGAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCA 1500  
QY ACCAGAGAAATGAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAAT 1500  
Db 1441 ACCAGAGAAATGAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCA 1500  
QY ACCAGAGAAATGAGAAATGAGACACTTCATCACTCAATCAATCAATCAATCAAT 1500  
Db 1501 TCCAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
QY TCCAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
Db 1501 TCCAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
QY TCCAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
Db 1561 ATCTTAAAGCTGAGTAAAGCTGAGTAAAGCTGAGTAAAGCTGAGTAAAGCTGAG 1620  
QY ATCTTAAAGCTGAGTAAAGCTGAGTAAAGCTGAGTAAAGCTGAGTAAAGCTGAG 1620  
Db 1621 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
QY GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
Db 1621 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
QY GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
Db 1681 TTCAACTTCAACCAAGTAAAGCTGCTTCAATTCACAGAACCCAAACGATCGAG 1740  
QY TTCAACTTCAACCAAGTAAAGCTGCTTCAATTCACAGAACCCAAACGATCGAG 1740  
Db 1741 AAGCACTTCTAGTAAAGCTGCTTCAATTCACAGAACCCAAACGATCGAGTCTAT 1800  
QY AAGCACTTCTAGTAAAGCTGCTTCAATTCACAGAACCCAAACGATCGAGTCTAT 1800  
Db 1741 AAGCACTTCTAGTAAAGCTGCTTCAATTCACAGAACCCAAACGATCGAGTCTAT 1800  
QY AAGCACTTCTAGTAAAGCTGCTTCAATTCACAGAACCCAAACGATCGAGTCTAT 1800  
Db 1801 CTCCTCAGGAG 1860  
QY CTCCTCAGGAG 1860  
Db 1801 CTCCTCAGGAG 1860  
QY CTCCTCAGGAG 1860  
Db 1861 TACCTAGAGCTGAG 1920  
QY TACCTAGAGCTGAG 1920  
Db 1861 TACCTAGAGCTGAG 1920  
QY TACCTAGAGCTGAG 1920  
Db 1921 GAGCAACAGTTCAG 1980  
QY GAGCAACAGTTCAG 1980  
Db 1921 GAGCAACAGTTCAG 1980  
QY GAGCAACAGTTCAG 1980  
Db 1981 AACCTGCTTGTGAGTCTTATCTCATGATTAACCTCTCGAGCTGTCTCCGAT 2040  
QY AACCTGCTTGTGAGTCTTATCTCATGATTAACCTCTCGAGCTGTCTCCGAT 2040

Db 1981 AACCTGCTTGTGAGTCTTATCTCATGATTAACCTCTCGAGCTGTCTCCGAT 2040  
QY AACCTGCTTGTGAGTCTTATCTCATGATTAACCTCTCGAGCTGTCTCCGAT 2040  
Db 2041 GCTGTGGCAAAATGCGGAGCTGCGGATTAAGTCAATGATGATGATGATGATGATG 2100  
QY GCTGTGGCAAAATGCGGAGCTGCGGATTAAGTCAATGATGATGATGATGATGATG 2100  
Db 2041 GCTGTGGCAAAATGCGGAGCTGCGGATTAAGTCAATGATGATGATGATGATGATG 2100  
QY GCTGTGGCAAAATGCGGAGCTGCGGATTAAGTCAATGATGATGATGATGATGATG 2100  
Db 2101 ATCAACCCAAAGCATTTGCTAAAGGGGAGGAGATTAATCTGAGAGAGTAAAGAG 2160  
QY ATCAACCCAAAGCATTTGCTAAAGGGGAGGAGATTAATCTGAGAGAGTAAAGAG 2160  
Db 2101 ATCAACCCAAAGCATTTGCTAAAGGGGAGGAGATTAATCTGAGAGAGTAAAGAG 2160  
QY ATCAACCCAAAGCATTTGCTAAAGGGGAGGAGATTAATCTGAGAGAGTAAAGAG 2160  
Db 2161 GAAACATTTGCTGCGGCTTCAACATTCAGTAAACCAAGTAAACCAAGTAAACCA 2220  
QY GAAACATTTGCTGCGGCTTCAACATTCAGTAAACCAAGTAAACCAAGTAAACCA 2220  
Db 2161 GAAACATTTGCTGCGGCTTCAACATTCAGTAAACCAAGTAAACCAAGTAAACCA 2220  
QY GAAACATTTGCTGCGGCTTCAACATTCAGTAAACCAAGTAAACCAAGTAAACCA 2220  
Db 2221 GCTGTGAGTAAATGAG 2280  
QY GCTGTGAGTAAATGAG 2280  
Db 2221 GCTGTGAGTAAATGAG 2280  
QY GCTGTGAGTAAATGAG 2280  
Db 2281 TTGGGATACACAG 2340  
QY TTGGGATACACAG 2340  
Db 2281 TTGGGATACACAG 2340  
QY TTGGGATACACAG 2340  
Db 2341 CTGAG 2400  
QY CTGAG 2400  
Db 2341 CTGAG 2400  
QY CTGAG 2400  
Db 2401 TCTCAGCTTTGAAAAAGAGAGATTTGAGGCTTGCATGAGAGATTTGAGGAGAT 2460  
QY TCTCAGCTTTGAAAAAGAGAGATTTGAGGCTTGCATGAGAGATTTGAGGAGAT 2460  
Db 2401 TCTCAGCTTTGAAAAAGAGAGATTTGAGGCTTGCATGAGAGATTTGAGGAGAT 2460  
QY TCTCAGCTTTGAAAAAGAGAGATTTGAGGCTTGCATGAGAGATTTGAGGAGAT 2460  
Db 2461 TCCAGCAAGCTGCTGAGATGATTTCTTCTGATGACAACTTGGCTGATGATGAG 2520  
QY TCCAGCAAGCTGCTGAGATGATTTCTTCTGATGACAACTTGGCTGATGATGAG 2520  
Db 2461 TCCAGCAAGCTGCTGAGATGATTTCTTCTGATGACAACTTGGCTGATGATGAG 2520  
QY TCCAGCAAGCTGCTGAGATGATTTCTTCTGATGACAACTTGGCTGATGATGAG 2520  
Db 2521 GTAG 2580  
QY GTAG 2580  
Db 2521 GTAG 2580  
QY GTAG 2580  
Db 2581 AGTAAATTCGAGAAATCAACCCCTTCTTGAATTAATTAATTAATTAATTAAT 2640  
QY AGTAAATTCGAGAAATCAACCCCTTCTTGAATTAATTAATTAATTAATTAAT 2640  
Db 2581 AGTAAATTCGAGAAATCAACCCCTTCTTGAATTAATTAATTAATTAATTAAT 2640  
QY AGTAAATTCGAGAAATCAACCCCTTCTTGAATTAATTAATTAATTAATTAAT 2640  
Db 2641 CTGGGAG 2700  
QY CTGGGAG 2700  
Db 2641 CTGGGAG 2700  
QY CTGGGAG 2700  
Db 2701 CTGGGAG 2760  
QY CTGGGAG 2760  
Db 2701 CTGGGAG 2760  
QY CTGGGAG 2760  
Db 2761 GACAACTTGTGAAG 2820  
QY GACAACTTGTGAAG 2820  
Db 2761 GACAACTTGTGAAG 2820  
QY GACAACTTGTGAAG 2820  
Db 2821 GCGCTGAGAGAGCTTCTTCACTTAATTTGTGATTTGCTGAGAGAGAGAGAGAG 2880  
QY GCGCTGAGAGAGCTTCTTCACTTAATTTGTGATTTGCTGAGAGAGAGAGAGAGAG 2880  
Db 2821 GCGCTGAGAGAGCTTCTTCACTTAATTTGTGATTTGCTGAGAGAGAGAGAGAG 2880  
QY GCGCTGAGAGAGCTTCTTCACTTAATTTGTGATTTGCTGAGAGAGAGAGAGAGAG 2880  
Db 2881 CACTGTTGGGAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
QY CACTGTTGGGAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
Db 2881 CACTGTTGGGAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
QY CACTGTTGGGAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
Db 2941 TACGGGAG 3000  
QY TACGGGAG 3000  
Db 2941 TACGGGAG 3000  
QY TACGGGAG 3000  
Db 3001 TTCTTTGTCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
QY TTCTTTGTCAATGATGAG 3060  
Db 3001 TTCTTTGTCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
QY TTCTTTGTCAATGATGAG 3060  
Db 3061 AATTCTGCTTCCAG 3120  
QY AATTCTGCTTCCAG 3120







Dh 541 ATGTACTGTGATGTGAGGCATTCTTTGTTCTTGCTTATGCAATCCGAGTGTCTACA 600  
Qy 601 GAAGAGAACCCCAATATGATCTGTACTCGGGGTGCGTGTGCTGTGCTATC 660  
Dh 601 GAAGAGAACCCCAATATGATCTGTACTCGGGGTGCGTGTGCTGTGCTATC 660  
Qy 661 ATAACTGGCTGTCTCTCTATTTATCAAGACAAAAGCTCCAGATCATGAACTCTTC 720  
Dh 661 ATAACTGGCTGTCTCTCTATTTATCAAGACAAAAGCTCCAGATCATGAACTCTTC 720  
Qy 721 AAGAACTGTCTCTCTCAAGCAAGCCCTGTGATTTGAAATGGAAGAAATGAGATCAAC 780  
Dh 721 AAGAACTGTCTCTCTCAAGCAAGCCCTGTGATTTGAAATGGAAGAAATGAGATCAAC 780  
Qy 781 GCAGAGAGATGTCTGT 840  
Dh 781 GCAGAGAGATGTCTGT 840  
Qy 841 GATCTCAGAAATCATATCTGCAAAATGTGTGCAAGGTGATTAATCTCTCACTGATGTGAA 900  
Dh 841 GATCTCAGAAATCATATCTGCAAAATGTGTGCAAGGTGATTAATCTCTCACTGATGTGAA 900  
Qy 901 TCAGAACCCCAAGATCTGGTCTCCCGGATTTTCAAAAGCAAGCCCTGTGAGACAAAGAAC 960  
Dh 901 TCAGAACCCCAAGATCTGGTCTCCCGGATTTTCAAAAGCAAGCCCTGTGAGACAAAGAAC 960  
Qy 961 ATTGCTTCTTCTCAACCACTGTGTGAAAGAACTGACGTGTGTGTGTGTGTGTGTGTGT 1020  
Dh 961 ATTGCTTCTTCTCAACCACTGTGTGAAAGAACTGACGTGTGTGTGTGTGTGTGTGTGT 1020  
Qy 1021 GGGATGTGACCTGT 1080  
Dh 1021 GGGATGTGACCTGT 1080  
Qy 1081 ACCCCCATTTGTGAAAGAAATGTGAGCACTTTCATCACTGATCAACGGGTGTGTGTGTGT 1140  
Dh 1081 ACCCCCATTTGTGAAAGAAATGTGAGCACTTTCATCACTGATCAACGGGTGTGTGTGTGT 1140  
Qy 1141 CTGGGGGTGTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
Dh 1141 CTGGGGGTGTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
Qy 1201 ATCTTCTCTTGT 1260  
Dh 1201 ATCTTCTCTTGT 1260  
Qy 1261 GTATGTCTGACCTCACTGTGCAAGGCGATGTGCGAGAAAGTTGTGTGTGTGTGTGTGT 1320  
Dh 1261 GTATGTCTGACCTCACTGTGCAAGGCGATGTGCGAGAAAGTTGTGTGTGTGTGTGTGT 1320  
Qy 1321 GAAGCTGTGAGACCTTGGGGGTGCACATCCACATCTGCTCCGACAAAGCTGGAACTCTG 1380  
Dh 1321 GAAGCTGTGAGACCTTGGGGGTGCACATCCACATCTGCTCCGACAAAGCTGGAACTCTG 1380  
Qy 1381 ACTCAGAACCCGATGACAGTGTGTGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440  
Dh 1381 ACTCAGAACCCGATGACAGTGTGTGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440  
Qy 1441 ACCACAGAGAAATGAGAGTGGGGGTCTCTTGTGACAAAGCGTCAAGCGTGTGTGTGTGT 1500  
Dh 1441 ACCACAGAGAAATGAGAGTGGGGGTCTCTTGTGACAAAGCGTCAAGCGTGTGTGTGTGT 1500  
Qy 1501 TCCGAAATTTCTGT 1560  
Dh 1501 TCCGAAATTTCTGT 1560  
Qy 1561 ATCTTAAAGGT 1620  
Dh 1561 ATCTTAAAGGT 1620  
Qy 1621 GTCTGT 1680  
Dh 1621 GTCTGT 1680

Qy 1681 TTCAATCCCAACCAATGATACAGTCTTCCATTTCAACAAAGAACCAAGCAATGGAGCT 1740  
Dh 1681 TTCAATCCCAACCAATGATACAGTCTTCCATTTCAACAAAGAACCAAGCAATGGAGCT 1740  
Qy 1741 AAGCACTCTGT 1800  
Dh 1741 AAGCACTCTGT 1800  
Qy 1801 CTCTCTCAGGCAAGGAGAGAGCCCTGTGACGAAGCTGTGAAGAGAGAGAGAGAGAGAGAG 1860  
Dh 1801 CTCTCTCAGGCAAGGAGAGAGCCCTGTGACGAAGCTGTGAAGAGAGAGAGAGAGAGAGAG 1860  
Qy 1861 TACCTGAGCTGT 1920  
Dh 1861 TACCTGAGCTGT 1920  
Qy 1921 GACGAACAGTTTCCCGAAGGCTTCCAGTTGTGACACTGATGAAGTCAATTTCCCGTGTGT 1980  
Dh 1921 GACGAACAGTTTCCCGAAGGCTTCCAGTTGTGACACTGATGAAGTCAATTTCCCGTGTGT 1980  
Qy 1981 AACCTGTCTGT 2040  
Dh 1981 AACCTGTCTGT 2040  
Qy 2041 GCTGTGGCAAAATGTGCGACAGCGCTGTGAGATTTAAGTCAATGTGTCAAGAGACATGCA 2100  
Dh 2041 GCTGTGGCAAAATGTGCGACAGCGCTGTGAGATTTAAGTCAATGTGTGTCAAGAGACATGCA 2100  
Qy 2101 ATCAACGCAAAAGCATGT 2160  
Dh 2101 ATCAACGCAAAAGCATGT 2160  
Qy 2161 GAAGACATTTGT 2220  
Dh 2161 GAAGACATTTGT 2220  
Qy 2221 GCTGT 2280  
Dh 2221 GCTGT 2280  
Qy 2281 TTGGGGT 2340  
Dh 2281 TTGGGGT 2340  
Qy 2341 GTGAGAGCTGT 2400  
Dh 2341 GTGAGAGCTGT 2400  
Qy 2401 TCTCCAGCTTTGAAAAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460  
Dh 2401 TCTCCAGCTTTGAAAAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460  
Qy 2461 TCCAAAGCAAGCTGT 2520  
Dh 2461 TCCAAAGCAAGCTGT 2520  
Qy 2521 GTAGAAGAAAGT 2580  
Dh 2521 GTAGAAGAAAGT 2580  
Qy 2581 AGTAACATTTCCGAAATCAACCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640  
Dh 2581 AGTAACATTTCCGAAATCAACCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640  
Qy 2641 CTGGGCACTGT 2700  
Dh 2641 CTGGGCACTGT 2700  
Qy 2701 CTGGGCTATGAACAGGCTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760  
Dh 2701 CTGGGCTATGAACAGGCTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760







QY 421 GGCTTAACCCCGCAGGCGCGTGAATCTGGCTGGGATGGCCCCCAAGCCCTCAG 480  
DB 421 GGCTTAACCCCGCAGGCGCGTGAATCTGGCTGGGATGGCCCCCAAGCCCTCAG 480  
QY 481 CCCCCCTCCACTACTCCCGAGTGGGTCAAAATTCGTGGGAGCTGTTGGGTGCTTCC 540  
DB 481 CCCCCCTCCACTACTCCCGAGTGGGTCAAAATTCGTGGGAGCTGTTGGGTGCTTCC 540  
QY 541 ATGTACTGTGATTTGAGAGCCATCTTTGTTCTTGAGCTTAATGGCATCCGAATGTACA 600  
DB 541 ATGTACTGTGATTTGAGAGCCATCTTTGTTCTTGAGCTTAATGGCATCCGAATGTACA 600  
QY 601 GAAAGAGAACCCAAATGATGATCTGACTCGGGGTGGGTGCTGTCTGTCTGATC 660  
DB 601 GAAAGAGAACCCAAATGATGATCTGACTCGGGGTGGGTGCTGTCTGTCTGATC 660  
QY 661 ATAACTGGCTGTTCTCCTATTAATCAAGAACAAAAGCTCCAAAGTCAATGATCTTTC 720  
DB 661 ATAACTGGCTGTTCTCCTATTAATCAAGAACAAAAGCTCCAAAGTCAATGATCTTTC 720  
QY 721 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
DB 721 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
QY 781 GCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 781 GCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 841 GATCTCAGAAATCATATCTGCAAAATGATGATGATGATGATGATGATGATGATGAT 900  
DB 841 GATCTCAGAAATCATATCTGCAAAATGATGATGATGATGATGATGATGATGATGAT 900  
QY 901 TCAGAACCCCGACACTGGTCCCGGATTTCAAAACGAGAACCCCTTGAGACAAAGAAC 960  
DB 901 TCAGAACCCCGACACTGGTCCCGGATTTCAAAACGAGAACCCCTTGAGACAAAGAAC 960  
QY 961 ATTGCTCTCTCTCAACCAACTGTGTGTAAGAACTGCAAGTGGCATCGTTGTGATCACT 1020  
DB 961 ATTGCTCTCTCTCAACCAACTGTGTGTAAGAACTGCAAGTGGCATCGTTGTGATCACT 1020  
QY 1021 GGGGATTCGACCGGTATGAGGAGATCGCCACCTTGTCTTGGGTGGAAGGCGCTTG 1080  
DB 1021 GGGGATTCGACCGGTATGAGGAGATCGCCACCTTGTCTTGGGTGGAAGGCGCTTG 1080  
QY 1081 ACCCCATTTGCTGAAGAAATCGAGCACTTCATCCACTCATCAAGGATGAGCGGTGTC 1140  
DB 1081 ACCCCATTTGCTGAAGAAATCGAGCACTTCATCCACTCATCAAGGATGAGCGGTGTC 1140  
QY 1141 CTGGGGGTGCTTTCTTCAATCTCTCTGTGATCTTGAATCACTGGCTCGAGCTGTC 1200  
DB 1141 CTGGGGGTGCTTTCTTCAATCTCTCTGTGATCTTGAATCACTGGCTCGAGCTGTC 1200  
QY 1201 ATCTTCCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
DB 1201 ATCTTCCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
QY 1261 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
DB 1261 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
QY 1321 GAAAGCTGAGAGACCTTTGGGTGATCATCATCATCATCATCATCATCATCATCATCAT 1380  
DB 1321 GAAAGCTGAGAGACCTTTGGGTGATCATCATCATCATCATCATCATCATCATCATCAT 1380  
QY 1381 ACTCAGAACCGGATGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
DB 1381 ACTCAGAACCGGATGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
QY 1441 ACCACAGAGATGAGTGGGTGCTTCTTGAACAAGACGTCAGACCACTGGTGCCTG 1500  
DB 1441 ACCACAGAGATGAGTGGGTGCTTCTTGAACAAGACGTCAGACCACTGGTGCCTG 1500  
QY 1501 TCAGAAATTCCTGCTCTGTAACAGGAGAGTGTTCAGGCTAAACCAAGAAACCTGCT 1560

DB 1501 TCAGAAATTCCTGCTCTGTAACAGGAGAGTGTTCAGGCTAAACCAAGAAACCTGCT 1560  
QY 1561 ATCTTAAGCTGAGATGACGGAGATGCTTCCAGATGGGCTTTTAAATGATGATGATGAT 1620  
DB 1561 ATCTTAAGCTGAGATGACGGAGATGCTTCCAGATGGGCTTTTAAATGATGATGATGAT 1620  
QY 1621 GTCGAGTGGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
DB 1621 GTCGAGTGGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
QY 1681 TTCAACTCCACCAACAACTAGTACAGCTTCAATTCACAAAGACCAAGCATGAGACT 1740  
DB 1681 TTCAACTCCACCAACAACTAGTACAGCTTCAATTCACAAAGACCAAGCATGAGACT 1740  
QY 1741 AAGCACTGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
DB 1741 AAGCACTGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
QY 1801 CTCTCCACGGCAAGAGACAGCCCTGAGACGAGAGCTGAAAGAGCCCTTTCAGAAATGCC 1860  
DB 1801 CTCTCCACGGCAAGAGACAGCCCTGAGACGAGAGCTGAAAGAGCCCTTTCAGAAATGCC 1860  
QY 1861 TACCTAAGCTGGGGGCTTGGAGAGCTGCTGATGATGATGATGATGATGATGATGATGAT 1920  
DB 1861 TACCTAAGCTGGGGGCTTGGAGAGCTGCTGATGATGATGATGATGATGATGATGATGAT 1920  
QY 1921 GACGAACGTTTCCGAAAGCTTCCAGTTTGACACTGATGAAGTCAATTTCCCGTGTAT 1980  
DB 1921 GACGAACGTTTCCGAAAGCTTCCAGTTTGACACTGATGAAGTCAATTTCCCGTGTAT 1980  
QY 1981 AACCTGTGCTGAGGCTTATATCTTCATGATTTGACCCCTCTGAGCTGTCTCCCGAT 2040  
DB 1981 AACCTGTGCTGAGGCTTATATCTTCATGATTTGACCCCTCTGAGCTGTCTCCCGAT 2040  
QY 2041 GCTGTGGGAAATGCGCGAGCGCTGGGATTAAGTATGATGATGATGATGATGATGATGATGAT 2100  
DB 2041 GCTGTGGGAAATGCGCGAGCGCTGGGATTAAGTATGATGATGATGATGATGATGATGATGAT 2100  
QY 2101 ATCAGACCCAAAGCCATTTGTAAGGGGATGAGCATTTATCTCAAGAAATGACGAGCCGTG 2160  
DB 2101 ATCAGACCCAAAGCCATTTGTAAGGGGATGAGCATTTATCTCAAGAAATGACGAGCCGTG 2160  
QY 2161 GAAAGACTGCTGCTCCGCTCAAACTTCCAGTGAACCAAGTGAACCCCGAGATGGCAAG 2220  
DB 2161 GAAAGACTGCTGCTCCGCTCAAACTTCCAGTGAACCAAGTGAACCCCGAGATGGCAAG 2220  
QY 2221 GCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
DB 2221 GCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
QY 2281 TTGCGGTACCAACCGGAGATGCTTGTGTAAGACTTTCCTTAACAGAAAGCTCATATT 2340  
DB 2281 TTGCGGTACCAACCGGAGATGCTTGTGTAAGACTTTCCTTAACAGAAAGCTCATATT 2340  
QY 2341 GTGAGAGCTGCGAGGCGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
DB 2341 GTGAGAGCTGCGAGGCGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
QY 2401 TCTCCAGCTTTGAAAAAGGAGATTTGGGATGCGATGGGATTTGGCTCGGATGTG 2460  
DB 2401 TCTCCAGCTTTGAAAAAGGAGATTTGGGATGCGATGGGATTTGGCTCGGATGTG 2460  
QY 2461 TCCAGAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520  
DB 2461 TCCAGAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520  
QY 2521 GTAGAAAGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
DB 2521 GTAGAAAGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
QY 2581 AGTAACATTCGGAAATCACCCCTTCTGATATTTATTTATGTAACCATTTCCCTGCC 2640



Dh	2581	AGTAAACATCGGAAACCAACACCCCTTTTGTATATATTATTATGCAACCACTGCAAGTCC	2640
Cy	2441	CTGCGGACATCGTACATATCTCTGTATTTGATTTGCTGACATGCTTCCGCCATCTCT	2700
Dh	2741	TCTTCAACGCGACATCTCTGATTAATCTTCTGCACTGCAATATCTCCCGCATCTCT	2760
Cy	2761	CTGGGTATATGAAATAGCTTGAAGCTATCATATGAAAGAGAGAGAGAAATCCGAAAG	2760
Dh	2701	CTGGTATATGAAAGCTTGAAGCTATCATATGAAAGAGAGAGAGAAATCCGAAAG	2760
Cy	2761	GACAAACTCTTGAAGCAAGCTTTGATAGCATGCTGATGACACGATGAGATGATGAG	2820
Dh	2761	GACAAACTCTTGAAGCAAGCTTTGATAGCATGCTGATGAGCATGAGATGATGAG	2820
Cy	2821	GACCTGGAAGAGATCTTATATTTTGTACTTCTGATGAGAAAGGTTCTGCTGCTTT	2880
Dh	2821	GACCTGGAAGAGATCTTATATTTTGTACTTCTGATGAGAAAGGTTCTGCTGCTTT	2880
Cy	2881	TACTTTTGTGATATCTGATGAGAAATCTGGAATGACCGTGTATATGATGATGATG	2940
Dh	2881	TACTTTTGTGATATCTGATGAGAAATCTGGAATGACCGTGTATATGATGATGATG	2940
Cy	2941	TACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3000
Dh	2941	TACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3000
Cy	3001	TTCTTTTGTATATCTGATGATGATGATGATGATGATGATGATGATGATGATG	3060
Dh	3001	TTCTTTTGTATATCTGATGATGATGATGATGATGATGATGATGATGATGATG	3060
Cy	3061	AAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3120
Dh	3061	AAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3120
Cy	3121	ACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3180
Dh	3121	ACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3180
Cy	3181	CTGCTGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG	3240
Dh	3181	CTGCTGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG	3240
Cy	3241	GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3300
Dh	3241	GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3300
Cy	3301	TACTACTAGCTATGATGATGATGATGATGATGATGATGATGATGATGATG	3360
Dh	3301	TACTACTAGCTATGATGATGATGATGATGATGATGATGATGATGATGATG	3360
Cy	3361	TACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3420
Dh	3361	TACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3420
Cy	3421	GCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3480
Dh	3421	GCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3480
Cy	3481	GACGATGATGATGATGATGATGATGATGATGATGATGATGATG	3540
Dh	3481	GACGATGATGATGATGATGATGATGATGATGATGATGATGATG	3540
Cy	3541	CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	3600
Dh	3541	CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	3600
Cy	3601	GCATTTTATAAATGAAATATGATGATGATGATGATGATGATGATGATG	3660
Dh	3601	GCATTTTATAAATGAAATATGATGATGATGATGATGATGATGATGATG	3660

[illegible]



QY 61 GCGCAGACGCGGCGCTCGGCTCCGGGGCGCGGCGCTCTCTTCTCGCGGG 120  
DB 61 GCGGAGAGGCGGCGCTCGGCTCGGGGCGCGGCGCTCTCTTCTCGCGGG 120  
QY 121 CAGCTCTTCTAGTCTCCAGCCACAGACCCCGGCGCGGGCGCGGCGCTGCT 180  
DB 121 CAGGCTCTTCTAGTCTCCAGCCACAGACCCCGGCGCGGGCGCGGCGCTGCT 180  
QY 181 CTCTCTTCTAGTCTCCAGCCACAGACCCCGGCGCGGGCGCGGCGCTGCT 240  
DB 181 CTCTCTTCTAGTCTCCAGCCACAGACCCCGGCGCGGGCGCGGCGCTGCT 240  
QY 241 GGGAAAGGGGGTGGAGCAGACAGATAGAGCCCGAGCTGTATCAAGAACATGGGAGCAAG 300  
DB 241 GGGAAAGGGGGTGGAGCAGACAGATAGAGCCCGAGCTGTATCAAGAACATGGGAGCAAG 300  
QY 301 AAGAGCAAGAAAGCGAAGAAAGGAGGACATGACGAACTCAAGAGGAGTGTCTATG 360  
DB 301 AAGAGCAAGAAAGCGAAGAAAGGAGGACATGAGAACCTCAAGAGGAGTGTCTATG 360  
QY 361 GACGACCATAAACTCAGCTGATGAACTCCATCGTAAATACGGAACAGATTTGAGCGA 420  
DB 361 GACGACCATAAACTCAGCTGATGAACTCCATCGTAAATACGGAACAGATTTGAGCGA 420  
QY 421 GGGCTAAACCGCGAAGGGCGGCTGAGATCGGCTCGGGATGGCCCGAAGCGCTCAG 480  
DB 421 GGGCTAAACCGCGAAGGGCGGCTGAGATCGGCTCGGGATGGCCCGAAGCGCTCAG 480  
QY 481 CCGCTCCCACTACTCCGAGTGGGTCAAAATCTGTGCGAGCTGTTGCGTGGCTTCTCC 540  
DB 481 CCGCTCCCACTACTCCGAGTGGGTCAAAATCTGTGCGAGCTGTTGCGTGGCTTCTCC 540  
QY 541 ATGTTACTGTGATTTGAGCCATCTTGTGTTCTTGGCTTATGGCATCCGAGATGCTACA 600  
DB 541 ATGTTACTGTGATTTGAGCCATCTTGTGTTCTTGGCTTATGGCATCCGAGATGCTACA 600  
QY 601 GAAGAGAAACCAACCAATATGATCTGTACCTCGGGGTGCTGTCTGTGCTGCTATC 660  
DB 601 GAAGAGAAACCAACCAATATGATCTGTACCTCGGGGTGCTGTCTGTGCTGCTATC 660  
QY 661 ATATCTGCTGTTTCTCTCTATTTATCAAGAAAGCAAAAGCTCCAAAGATCATGGAATCCTTC 720  
DB 661 ATATCTGCTGTTTCTCTCTATTTATCAAGAAAGCAAAAGCTCCAAAGATCATGGAATCCTTC 720  
QY 721 AAGAACATGATGCTCCCTCAGCAAGCCCTCGATTTCAAAATGAGAGAAAGATGAGATCAAC 780  
DB 721 AAGAACATGATGCTCCCTCAGCAAGCCCTCGATTTCAAAATGAGAGAAAGATGAGATCAAC 780  
QY 781 GAGAGAGATGCTGCTGTTGATCTGTGTGAGGTGAAGGGCGAGACCGAATCCCTGCT 840  
DB 781 GAGAGAGATGCTGCTGTTGATCTGTGTGAGGTGAAGGGCGAGACCGAATCCCTGCT 840  
QY 841 GATCTAGAAATCAATCTGTGAAATGGCTGCAAGTGAATACTCCTCACTCACTGCTGAA 900  
DB 841 GATCTAGAAATCAATCTGTGAAATGGCTGCAAGTGAATACTCCTCACTCACTGCTGAA 900  
QY 901 TGAAGACCCCGAGACTCGTCCCGGATTTTCAAAACGAGAAACCCCTTGAGAGCAAGAAC 960  
DB 901 TGAAGACCCCGAGACTCGTCCCGGATTTTCAAAACGAGAAACCCCTTGAGAGCAAGAAC 960  
QY 961 ATTGCTCTTCTTCAACCAACTGTGTTGAAGAACTGACATGCTGTGTATCACT 1020  
DB 961 ATTGCTCTTCTTCAACCAACTGTGTTGAAGAACTGACATGCTGTGTATCACT 1020  
QY 1021 GGGGATGACCGCTGATGGCAGAGATGCGACCTTCTCTGCTGTAAGGCGGCTG 1080  
DB 1021 GGGGATGACCGCTGATGGCAGAGATGCGACCTTCTCTGCTGTAAGGCGGCTG 1080  
QY 1081 ACCGCTCTTGTGAAGAAATGAGACATTTATCTCACTCACTGCTGTGCTGCTG 1140  
DB 1081 ACCGCTCTTGTGAAGAAATGAGACATTTATCTCACTCACTGCTGTGCTGCTGCTG 1140

QY 1141 CTGGGGGTCTCTTCTTCTCATCTCTCTGSAICTTGAATACCTGGCTCGAGGATGTG 1200  
DB 1141 CTGGGGGTCTCTTCTTCTCATCTCTCTGATCTTGAATACCTGGCTCGAGGATGTG 1200  
QY 1201 ATCTTCTCATTTGATGATCATGTAGCAAGCGTCCGAAAGTTTGTGTGCGACCTGACG 1260  
DB 1201 ATCTTCTCATTTGATGATCATGTAGCAAGCGTCCGAAAGTTTGTGTGCGACCTGACG 1260  
QY 1261 GTATGTGACGCTCACTGCGCAAGCGCATGCGAGGAATGCTGCTGTGAAGAACTG 1320  
DB 1261 GTATGTGACGCTCACTGCGCAAGCGCATGCGAGGAATGCTGCTGTGAAGAACTG 1320  
QY 1321 GAAGCTGTGAGACCTTGGGATCCACATCCACATCTGTCTGACAAAGATGAACTTCTG 1380  
DB 1321 GAAGCTGTGAGACCTTGGGATCCACATCCACATCTGTCTGACAAAGATGAACTTCTG 1380  
QY 1381 ACTCGAAGCCGATGACATGTGCTCAATGTGGTTTGAACATTCATGAAGCTGAC 1440  
DB 1381 ACTCGAAGCCGATGACATGTGCTCAATGTGGTTTGAACATTCATGAAGCTGAC 1440  
QY 1441 ACCAGAGAAATCAGAGTGGGCTCTCTTGAACAAGCTGACCACTGTTGCTCTG 1500  
DB 1441 ACCAGAGAAATCAGAGTGGGCTCTCTTGAACAAGCTGACCACTGTTGCTCTG 1500  
QY 1501 TCCAGAAATGCTGTCTCTGTAAACAGGCGATGTTCAAGCTTAAACCAAGAAACCTGCT 1560  
DB 1501 TCCAGAAATGCTGTCTCTGTAAACAGGCGATGTTCAAGCTTAAACCAAGAAACCTGCT 1560  
QY 1561 ATCTTAAAGCGAGTACAGGAGATGCTCCGAGTGGGCTCTTAAAGTGCATGAG 1620  
DB 1561 ATCTTAAAGCGAGTACAGGAGATGCTCCGAGTGGGCTCTTAAAGTGCATGAG 1620  
QY 1621 GTCTGTGTGCTCCGTGAATGAGATGAGGAGAAAGTACACCAAGATGAGGATTCCT 1680  
DB 1621 GTCTGTGTGCTCCGTGAATGAGATGAGGAGAAAGTACACCAAGATGAGGATTCCT 1680  
QY 1681 TTCAACTCCACCAAGTAACCAAGCTCTCATTCACAAAGAACCAAGCATGAGACCT 1740  
DB 1681 TTCAACTCCACCAAGTAACCAAGCTCTCATTCACAAAGAACCAAGCATGAGACCT 1740  
QY 1741 AAGCACTGTAGTATGAAGGGCGGCCCAAGAAAGATCCGTGAGCCGATGACTATC 1800  
DB 1741 AAGCACTGTAGTATGAAGGGCGGCCCAAGAAAGATCCGTGAGCCGATGACTATC 1800  
QY 1801 CTCTTCAAGGCAAGACAGCCCTTGAGAGAGTGAAGAGCGCTTCAAGATGCC 1860  
DB 1801 CTCTTCAAGGCAAGACAGCCCTTGAGAGAGTGAAGAGCGCTTCAAGATGCC 1860  
QY 1861 TACCTAGAGCTGGGGGCTTGAAGAGCGTGTAGTGTCTGCACTCTTCTGCT 1920  
DB 1861 TACCTAGAGCTGGGGGCTTGAAGAGCGTGTAGTGTCTGCACTCTTCTGCT 1920  
QY 1921 GAGCAAGTTTCCCGAAGGCTTCAGTTGACATGATGAAGTCAATTTCCCGTGGAT 1980  
DB 1921 GAGCAAGTTTCCCGAAGGCTTCAGTTGACATGATGAAGTCAATTTCCCGTGGAT 1980  
QY 1981 AACCTGCTTGTGGGCTTATCTCATGATTAACCTCCTGAGCTGCTGCCGAT 2040  
DB 1981 AACCTGCTTGTGGGCTTATCTCATGATTAACCTCCTGAGCTGCTGCCGAT 2040  
QY 2041 GCTGTGGGCAAAATGCGCAGCGCTGGATTTAAGGTCAATGCTGACAGGACCATCA 2100  
DB 2041 GCTGTGGGCAAAATGCGCAGCGCTGGATTTAAGGTCAATGCTGACAGGACCATCA 2100  
QY 2101 ATCAAGACCAAGCATTTGCTAAGGGGTGGGCAATTCAGAAAGTAAAGAGCGCTG 2160  
DB 2101 ATCAAGACCAAGCATTTGCTAAGGGGTGGGCAATTCAGAAAGTAAAGAGCGCTG 2160  
QY 2161 GAAGACATTTGTGCGCGCTCAACATTCAGATGAAGCAAGTGAAGCCCAAGATGCCAAG 2220  
DB 2161 GAAGACATTTGTGCGCGCTCAACATTCAGATGAAGCAAGTGAAGCCCAAGATGCCAAG 2220  
QY 2221 GCTGTGTAGTACATGGCAGTGAATTTAAGGACATGACCTTGAAGGAGCTGGATGACATT 2280







































Best Local Similarity 86.4%; Pred. No. 0;  
Matches 2496; Conservative 0; Mismatches 320; Indels 73; Gaps 12;

QY 760 GGAAGAAAGATGAGCATCAACGAGAGATGTCGCTGATGATCTGTGGAGGTGAG 819  
Db 6246 GGTAGACAAATGACCATTAATGCGAGAGATGTTGTGGGATCTGTGGAACTAAAA 6187  
QY 820 GCGGAGACCGAATCCCGCTGATCTCAGAAATCATATCTGCAAAATGCGTG--CAAGTGG 877  
Db 6186 GGAAGAGACCGAATCTCTGCTGACCTCGAATCATATCTGCAAAATGCGTGACAGGTGGA 6127  
QY 878 ATAACTCTCACTCACTGTGTAATAGAACCCGAGACTCGTCCCGGATTTCAAAAG 937  
Db 6126 TAACTCTCTGCTCACTGTGTAATAGAACCCGAGACTGAGTCTCAATTTCACAAAATG 6067  
QY 938 AGAACCCCTTGGAGACGAGAAACATTTGCTTCTCAACCAACTGTGTGAAGAACTG 997  
Db 6066 AAAACCCCTTGGAGACGAGAAACATTTGCTTCTTTCACCAA-TGTGTTGAAGGACCG 6008  
QY 998 CACGTGCGATCGTTGTGTACACTGAGGATCGACCGGTATGGGCAAGATCGCACCTTG 1057  
Db 6007 CACGTGATATAG-TGTCTAACACTGGGGAATCGCACTGTATGGAAAGAAATTCACACCTTG 5949  
QY 1058 CTTCTGGGCTGGAAAGCGGCGCTGACCCCATTTGCTGAAGAAATCGAGCACTTCATCGAC 1117  
Db 5948 CTTCTGGGCTGGAAAGAGGCGCAGACCCCATTTGCTGAGAAATTTGAACATTTATCGACA 5889  
QY 1118 TCATACCGGCTGGCGGCTGTTCTGGGGGTCTTTCTTCATCTCTCTGTATCTTG 1177  
Db 5888 TCATACCGGCTGGATGTTGTTCTGGGGGTGTTCTTCATCTCTCTGTATCTTG 5829  
QY 1178 AGTAACTCTG--CTTGAAGCTGTATCTTCTCATTTGATATCGTAG--CGAAGCTGCG 1235  
Db 5828 AGTAACTCTGAAATTTGAGGCTGTATCTTCTCATGATATCGTAG--CGAAGCTGCG 5769  
QY 1236 GAAAGTTTGTGAGCGAC--GTCACCGGTATGTTCTGACGCT--CACTGCGCAACCGCATGCG 1292  
Db 5768 GGAAGTTTGTGCGCGCGTGTGACGCGTGTGCTGACCTTACATTGCAAACTGCGATGGG 5709  
QY 1293 GA-GAAGAACTGCTGTGTAAGAAACCTGGAAGCTGTGAGACCTTGGGGTCCACATCC 1351  
Db 5708 AATGAAAAAACGCTTAGTGAAGAACTTGAAGCTGTGAGACCTTGGGGTCCACGTTCA 5649  
QY 1352 GCATCTGCTCCGACCAAGATCTGGAATCTGTAGAACCGATGACAGTGGCTCATATGT 1411  
Db 5648 GCATCTGCTCTGATATAAATCTGGAATCTGTAGAACCGATGACAGTGGCCCAATGT 5589  
QY 1412 GGTTCAGCAATCAATTCATGAAGCTGACACACAGAGAAATCGAGTGGGCTCTCTTG 1471  
Db 5588 GGTTCAGCAATCAATTCATGAAGCTGATACGACAGAGAAATCGAGTGGGCTCTCTTG 5529  
QY 1472 ACAAGCGTCAGCGACCTGGTGTGCTGTGCGAAGATGCTGTGCTGTGAACAGGGCAG 1531  
Db 5528 ACAAGCTTACGCTACCTGCTGCTGTGCTGTGCGAAGATGAGGCTCTTGTGAACAGGGCAG 5469  
QY 1532 TGTTCAGGCTAAACCAAGAAACCTGCTATCTTAAAGCGTGAAGTGGGAGATGCTT 1591  
Db 5468 TGTTCAGGCTAAACCAAGAAACCTACTATCTTAAAGCGGAGTGGCAGAGATGCTT 5409  
QY 1592 CCGAGTGGTGTCTTAAATGATGAGGCTGCTGTGTTCCGTAAAGAGATGAGG 1651  
Db 5408 CTGAGTCAGACTCTTAAATGATGAGGCTGCTGTGTTCCGTAAAGAGATGAGG 5349  
QY 1652 AGAAGTACCAAGATAGTGAATTCCTTCAACCTCCACCAACAGTACCAAGCTCTCA 1711  
Db 5348 AAGAGTACCAAGATAGTGAATTCCTTCAACCTCCACCAACAGTACCAAGTGTCTA 5289  
QY 1712 TTACAAAGAAACCAAGATGATGAGGCTTAAAGCACTGTGATGATGAAGGCGCCCGAG 1771  
Db 5288 TTACAAAGAAACCAAGATGATGAGGCTTAAAGCACTGTGATGATGAAGGCGCCCGAG 5229  
QY 1772 AAAGGATCTGAGCGGATGATGATGATCTCTCTCCAGCGCAAGAGAGCGCCCTGAGCG 1831

Db 5228 AAAGATCTTGAACCGTTGACGCTCTATCTCTCCAGCGCAAGAGAGCGCCCTGATG 5169  
QY 1832 AAGAGCTGAAGACCGCTTTGAGAAATGCTTACTTGAAGCTGGGGGCTTTGAGAGCGTG 1891  
Db 5168 AGAGCTGAAGACCGCTTTGAGAAATGCTTACTTGAAGCTGGGGGCTTTGAGAGCGTG 5109  
QY 1892 TGCTAGTTCCTGCGACCTCTCTGCTGCGAGGAAACAGTTTCCGAAAGGCTTCCAGTTTG 1951  
Db 5108 TCTAGGTTTCTGCGACCTCTTCTGCGAGATGAACAGTTTCTGAAGGCTTCCAGTTTG 5049  
QY 1952 ACATGATGAAGTCAATTTCCCGTGAATTAACCTCTGCTGCTGAGCTTATCTCCATGA 2011  
Db 5048 ACATGATGAAGTCAATTTCCCGTGAATTAACCTCTGCTGCTGAGCTTATCTCCATGA 4989  
QY 2012 TTGACCTCTCTGAGCTGTGTCCCGATGTGTGAGGAAATGCGGAGGCTGAGATTA 2071  
Db 4988 TTGACCTCTCTGAGGCGCGCTTCTGATGCGGTGAGGAAATGCGGAGGCTGAGATTA 4929  
QY 2072 AGCTCATATGATCAAGAGACCAATCCATCAACAGCAAGCCATTTGCTAAAGGAGTGG 2131  
Db 4928 AGCTCATATGATCAAGAGACCAATCCATCAACAGCTAAAGCTATTTGCAAAAGTGG 4869  
QY 2132 GCATTATCTCAGAAAGTAAAGAGACCGGTGAAAGACATTTGCTGCGCCCTCAACATTCAG 2191  
Db 4868 GCATCATCTCAGAAAGTAAAGAGACCGGTGAAAGACATTTGCTGCGCCCTCAACATTCAG 4809  
QY 2192 TGAACAGGTTGAACCCCAAGATGCAAGTGGTGTGATGATGATGATGATGATGATG 2251  
Db 4808 TGAACAGGTTGAACCCCAAGATGCAAGTGGTGTGATGATGATGATGATGATGATG 4749  
QY 2252 ACATGACCTCTGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATG 2311  
Db 4748 ACATGACCTCTGAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATG 4689  
QY 2312 GGAACCTCTCTCAAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATG 2371  
Db 4688 GGAACCTCTCTCAAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATG 4629  
QY 2372 TGGCTGTCAGAGGAGTGGTGTGATGATGATGATGATGATGATGATGATGATGATG 2431  
Db 4628 TGGCTGTCAGAGGAGTGGTGTGATGATGATGATGATGATGATGATGATGATGATG 4569  
QY 2432 TTGCGATGGGAGATTTGGTGTGATGATGATGATGATGATGATGATGATGATGATG 2491  
Db 4568 TTGCGATGGGAGATTTGGTGTGATGATGATGATGATGATGATGATGATGATGATG 4509  
QY 2492 ATGACAACTTTCCTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 2551  
Db 4508 ATGACAACTTTCCTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 4449  
QY 2552 TGAAGAAATTCATTTGCTTACACCTTAAAGATTAACATTTCCGGAATCAACCCCTTTGA 2611  
Db 4448 TGAAGAAATTCATTTGCTTACACCTTAAAGATTAACATTTCCGGAATCAACCCCTTTGA 4389  
QY 2612 TATTATTAATGCAACATTCACCTGCGCCCTGGGACCGTGAACATCTCTGATTTGACT 2671  
Db 4388 TATTATTAATGCAACATTCACCTGCGCCCTGGGACCGTGAACATCTCTGATTTGACT 4329  
QY 2672 TGGGCACTGACATGTTCCGCGCATCTCTGCGCTATGGAACAGGCTGAAAGTACATCA 2731  
Db 4328 TGGGCACTGACATGTTCCGCGCATCTCTGCGCTATGGAACAGGCTGAAAGTACATCA 4269  
QY 2732 TGAAGAGCGACCGCAAAATCCCAAAACGACAACTTGAAGAGAGGCTGATGACCA 2791  
Db 4268 TGAAGAGCGACCGCAAAATCCCAAAACGACAACTTGAAGAGAGGCTGATGACCA 4209  
QY 2792 TGGCTATGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2851  
Db 4208 TGGCTATGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4149  
QY 2852 TTTCTGCTGAGAAAGTTCCTGCTTCACTGATGATGATGATGATGATGATGATGATG 2911  
Db 4148 TTTCTGCTGAGAAAGTTCCTGCTTCACTGATGATGATGATGATGATGATGATGATG 4089















Db	788	TTCCAGAGCCCGACACTCGCTCTCCGACTGCACTCACGAGAAACCCCTTTAAGACTCGGAA	847
OY	960	CATTGCGCTTCTCTCAACCAACTGTGTTGAAGAACTGCACGCTGAGCATGTTGTATAC	1019
Db	848	CATACCGTCTCTTTCACAACTTTGTAGAAAGCAGCGCTCGGAGCGGTGTGGTGGCCAC	907
OY	1020	TGGGATGCGACCGTGAATGGCGAGATCGCCACCCTTGCTCTT36GCTGGAAAGCGGCT	1079
Db	908	GGGGCACCAGCATGTGATCGAGGCGGATCGCCACCCTGGGATCAGGGCTGGAAGTGGGCA	967
OY	1080	GACCCCAATTGCTGAAGAAATGAGACACTTCATCCACTCATCAGCGAGTGGCCGTGT	1139
Db	968	GAGCGCCATTCGCGATCGAATTAGACACTTCATCCAGCTCATACCGGCGGTGTCTT	1027
OY	1140	CCTGGGGGTGTCTTTCTTCATTCTCTCTGTATCTTGAATCCTTGAATCCTGCTCGAGCTGT	1199
Db	1028	CCTGGGTCCTCTCTCTTCATCTCTCTCCCTCATCTTCGGAATACCTGCTGGTTGAGCTGT	1087
OY	1200	CATCTTCTCATTTGATATATGTAGGCCAAGCGCCGAAAGTTGTGGCCACCGCTAC	1259
Db	1088	CATCTTCTCATTCGGCATATATGTGGCCAAATGTCCAGAGGCTCTGCTGGCCACTGTAC	1147
OY	1260	GGTATGTCTGACGCTCACTGCCAAGCGCATGGCGAGAAAGAACTGCTCTGTGAAGAACT	1319
Db	1148	TGTGTGTCTGACGCTGACCGCCCAAGCGCATGGCCGAAAGAACTGCTGTGTGAAGAACT	1207
OY	1320	GGAACCTGTGAGACTTGGGGGTCCCATCCACATCTGTCTCGCAAGAAGCTGGAACTCT	1379
Db	1208	GAGATGTGAAGAACTCTGAGCTGACAGTGCACATATGTGTCAAGATGAAGACAGAGACCT	1267
OY	1380	GACTCAGAAACCGATGACAGTGGGCTCACATGTGTGTTGACAAATCCATAGAGCTGA	1439
Db	1258	CACTCAGAAACCGCATGACAGTGGCCCAATGTGTGACAAACAGATCCACAGAGGCTGA	1327
OY	1440	CACCAAGAAATCAGAGTGGGGTCTCTCTTTACAAAGACTCAGCCACCCTGTGGCTCT	1499
Db	1328	CACCACTGAGGACAGTCAAGTCAAGGACCTCATTTGACAAAGTTGCCACACTGGGGTGGCT	1387
OY	1500	GTCCAAGAAAT - GCTGGTCTCTGTATACAGGGCAGTGTTCAGGCTAAACAAAGAAACCTG	1557
Db	1388	GTTTGACATCGCTGGGGTTTTGCAATCGCCCTGTCTTCAAGGGTGGTCAGGACATATC	1447
OY	1558	CCTATCCTTAAGCGTGCAGTAGCGGAGATGCTTCCGAGTGGGCGCTTTAAAGTCATC	1617
Db	1448	CCTGTGCTCAAGAGGATGTGGCTGGGGAGTCCCTCTGAGTCTCCCTGCTCAAGTGCATC	1507
OY	1618	GAGGTCTGCTGGGCTCCCGATGAGAGATGAGGAGAAAGTACCAAGATAGTGGAGATT	1677
Db	1508	GAGCTGTCTCTGGCTTCGTGAAGCTGATCGTGAACGCAACAAAGAAATGGCTGAGATT	1567
OY	1738	CCTAAGCACTGTACTGATGAAGGGGCGCCCAAGAAAGATCCTGAGCCGATGAGATTCT	1797
Db	1628	AAACGATACCTGTGTGATGAAGGGGTCGCCCGAGCGCATCTGAGCCGCTGCTCCACC	1687
OY	1798	ATCTCTCTCCACGGCAAGGACGACCCCTGAGCAGAAAGCTGAGAAGCGCTTTACAGAT	1857
Db	1688	ATCTCTCTCAAGGCAAGGACGACCTCTGAGCAGAGAAATGAAGGAGGCTTTCAGAAAT	1747
OY	1858	GCTTACTTGAAGTGTGGGGGCTCTTGAGAGCGTGTCTAGTTTCTCCACCTCCTCTG	1917
Db	1748	GCTTACTTGAAGTGTGGGTGGCGCTGGGGAGACGGCGTGTGGTTCCTCATTAATTAACCTG	1807
OY	1918	CTGACGAAACAGTTTCCCGAAGCTTCCAGTTTGACACTATGAAGTCAATTTCCCGGTG	1977
Db	1808	CGGAAAGAAAGTGTGGGCAAGGAGGATTTT3CTTTGATCTATAGAGGGAATTTACACAG	1867
OY	1978	GATTAAGTGTGGTGTGGGTCTTATCTTCATGAATGAACCTTCTCGAAGTGTGTGCTCC	2037

[illegible]







361 GACGACATAACTCAGCCCTGGATGAACCTCCATCGTAAATACGAAACAGATTCGAGCCGA 420



Db 361 GAGGACCAATACTCAGCCTGAGTGAACCTCATGTAAATACGAAACAGATTGAGCGCA 420  
QY 421 GGCTTAAGACCCGCAAGGGCCGCTGAGATCTCTGGCTCGGGATGAGCCGCAACGCGCTCAGC 480  
Db 421 GGCTTAAGACCCGCAAGGGCCGCTGAGATCTCTGGCTCGGGATGAGCCGCAACGCGCTCAGC 480  
QY 481 CCCCTCCCACTACTCCCGAGTGGGTCAAAATCTGTGCGAGCTGTTCGGTGGCTTCC 540  
Db 481 CCCCTCCCACTACTCCCGAGTGGGTCAAAATCTGTGCGAGCTGTTCGGTGGCTTCC 540  
QY 541 ATGTACTGTGAGTGGAGCCATTTCTTTGTTCTTGCTTATGAGATCCGAGT3CTACA 600  
Db 541 ATGTACTGTGAGTGGAGCCATTTCTTTGTTCTTGCTTATGAGATCCGAGT3CTACA 600  
QY 601 GAAGAGAAACCAACCAATGATGATCTGTACCTCGGGGCTGTGCTGTGCTGTGCTATC 660  
Db 601 GAAGAGAAACCAACCAATGATGATCTGTACCTCGGGGCTGTGCTGTGCTGTGCTATC 660  
QY 661 ATTAAGCTGCTGTTTCTCTTATTAATCAAGAAAGCAAAAAGCTCCAAATCATGAAATCCTTC 720  
Db 661 ATTAAGCTGCTGTTTCTCTTATTAATCAAGAAAGCAAAAAGCTCCAAATCATGAAATCCTTC 720  
QY 721 AAGAAATGCTCTCTCAGCAAC3CCTGATGATGCTGAAATG3AAGAAAGATGAGCATTAAC 780  
Db 721 AAGAAATGCTCTCTCAGCAAC3CCTGATGATGCTGAAATG3AAGAAAGATGAGCATTAAC 780  
QY 781 GGAGAGAGATGCTGCTGATGCTGATGCTGAGAGTGAAGGGCGAGAGCCGAATCCCTGCT 840  
Db 781 GGAGAGAGATGCTGCTGATGCTGATGCTGAGAGTGAAGGGCGAGAGCCGAATCCCTGCT 840  
QY 841 GATTTAAATATATATTTG3AAATG3TGAAGT3ATATATCTCTCACTCATCTG3TGA 900  
Db 841 GATTTAAATATATATTTG3AAATG3TGAAGT3ATATATCTCTCACTCATCTG3TGA 900  
QY 901 TTAGAACCCGAGACTGAGTCCCGGAGATTTCAACAAAGAGAAACCCCTTGGAGCAAGGAAC 960  
Db 901 TTAGAACCCGAGACTGAGTCCCGGAGATTTCAACAAAGAGAAACCCCTTGGAGCAAGGAAC 960  
QY 961 ATTGCTTCTTCAACCAACTGTGTGAAGAACTGCACTG3GATGCTGTGTACT 1020  
Db 961 ATTGCTTCTTCAACCAACTGTGTGAAGAACTGCACTG3GATGCTGTGTACT 1020  
QY 1021 G3G3AATG3AACCCTBANTG33A3A3ATG3CAACCTTGTCTTGGGCTGGAAGGCGCTG 1080  
Db 1021 G3G3AATG3AACCCTBANTG33A3A3ATG3CAACCTTGTCTTGGGCTGGAAGGCGCTG 1080  
QY 1081 ACCCCCAATG3TGAAGAAATG3AGAGACTTCACTCACTCAACG3GCTG3GCGCTGTC 1140  
Db 1081 ACCCCCAATG3TGAAGAAATG3AGAGACTTCACTCACTCAACG3GCTG3GCGCTGTC 1140  
QY 1141 CTGGGGGCTGCTTTCTTCACTTCTCTCTGATCTCTGAGTACACCTG3GCTGAGGCTGTC 1200  
Db 1141 CTGGGGGCTGCTTTCTTCACTTCTCTCTGATCTCTGAGTACACCTG3GCTGAGGCTGTC 1200  
QY 1201 APTTCTCATTTGATGATCATGATGAGCAACG3TGC3GAGAGTGTG3GCGCTCAGC 1260  
Db 1201 APTTCTCATTTGATGATCATGATGAGCAACG3TGC3GAGAGTGTG3GCGCTCAGC 1260  
QY 1261 GATATGCTCAAGCTCACTG3GCAAGCGCATG3GCAAGAACTG3GCGCTG3GAGAACTG 1320  
Db 1261 GATATGCTCAAGCTCACTG3GCAAGCGCATG3GCAAGAACTG3GCGCTG3GAGAACTG 1320  
QY 1321 GAAGCTGTGAGAGCTTTG3G3TGCATCTCACTCTGCTCCGCAAGACTG3GAAACTG 1380  
Db 1321 GAAGCTGTGAGAGCTTTG3G3TGCATCTCACTCTGCTCCGCAAGACTG3GAAACTG 1380  
QY 1381 ACTCAGAAACCGAGATGAGAGTGCATCATGATGATGATGATGATGATGATGATGATG 1440  
Db 1381 ACTCAGAAACCGAGATGAGAGTGCATCATGATGATGATGATGATGATGATGATGATG 1440  
QY 1441 AATAAGAAATATAGATG3G3TGTCTTGAAGAAAGTGCAGGCACTG3ATGCTG 1500  
Db 1441 AATAAGAAATATAGATG3G3TGTCTTGAAGAAAGTGCAGGCACTG3ATGCTG 1500

QY 1501 TCCAGATTTGCTGCTCTGTAACAGGCGAGTGTTCAGGCTTAACCAAGAAAACCTGCT 1560  
Db 1501 TCCAGATTTGCTGCTCTGTAACAGGCGAGTGTTCAGGCTTAACCAAGAAAACCTGCT 1560  
QY 1561 ATCTTAAGCGTGCAGTACCGGAGATGCTTCGAGTCCGCGCTCTTAAAGTGCATCGAG 1620  
Db 1561 ATCTTAAGCGTGCAGTACCGGAGATGCTTCGAGTCCGCGCTCTTAAAGTGCATCGAG 1620  
QY 1621 GTCTGCTGTGCTCCGCTGATGAGATGAGAGGAGAAATTAACCAAGTATGAGGATTTCC 1680  
Db 1621 GTCTGCTGTGCTCCGCTGATGAGATGAGAGGAGAAATTAACCAAGTATGAGGATTTCC 1680  
QY 1681 TTCATCTCACAACAAGTACAGCTCTCATCAACAACCAACCAAGCATGAGAGCT 1740  
Db 1681 TTCATCTCACAACAAGTACAGCTCTCATCAACAACCAACCAAGCATGAGAGCT 1740  
QY 1741 AAGCACTGCTAGTATGAGAGGCGCCCAAGAAAGATCCTGAGCCGATGCACTTATC 1800  
Db 1741 AAGCACTGCTAGTATGAGAGGCGCCCAAGAAAGATCCTGAGCCGATGCACTTATC 1800  
QY 1801 CTCCTCAGCGCAAGGAGAGAGCCCTGAGAGAAAGCTGAGAGAGCGCTTTCAGAAATG 1860  
Db 1801 CTCCTCAGCGCAAGGAGAGAGCCCTGAGAGAAAGCTGAGAGAGCGCTTTCAGAAATG 1860  
QY 1861 TACCTAGAGCTG3 1920  
Db 1861 TACCTAGAGCTG3 1920  
QY 1921 GAGGAAAGTTCCTCCCAAGGCTTCCAGTTGATGATGATGATGATGATGATGATGATGAT 1980  
Db 1921 GAGGAAAGTTCCTCCCAAGGCTTCCAGTTGATGATGATGATGATGATGATGATGATGAT 1980  
QY 1981 AACCTGCTGCTG3 2040  
Db 1981 AACCTGCTGCTG3 2040  
QY 2041 GCTGTG3G3CAATG3 2100  
Db 2041 GCTGTG3G3CAATG3 2100  
QY 2101 ATACAGCGCAAAAGCCATTTGTAAGG3G3G3G3G3G3G3G3G3G3G3G3G3G3G3G3G3G3 2160  
Db 2101 ATACAGCGCAAAAGCCATTTGTAAGG3G3G3G3G3G3G3G3G3G3G3G3G3G3G3G3G3G3 2160  
QY 2161 GAAGACATTTGCTGCGCGGCTCAACATTTCCAGTGAACAGGTGAACCCAGAGATG3GCAAG 2220  
Db 2161 GAAGACATTTGCTGCGCGGCTCAACATTTCCAGTGAACAGGTGAACCCAGAGATG3GCAAG 2220  
QY 2221 GCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
Db 2221 GCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
QY 2281 TTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
Db 2281 TTGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
QY 2341 GTGAGAGGCTGCGAGCGGCGAGG3 2400  
Db 2341 GTGAGAGGCTGCGAGCGGCGAGG3 2400  
QY 2401 TCTCCAGCTTTGAAAAG3 2460  
Db 2401 TCTCCAGCTTTGAAAAG3 2460  
QY 2461 TCCAGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520  
Db 2461 TCCAGCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520  
QY 2521 GTAGAGAAAGTGTGCTGATATTTGATTAACCTG3AAGAAATGATGCTTACACCTTAACA 2580  
Db 2521 GTAGAGAAAGTGTGCTGATATTTGATTAACCTG3AAGAAATGATGCTTACACCTTAACA 2580



QY 2581 AGTACATTCGGAATACCCCTTCTTGAATATTATTATTCGAACATTCCACTGCC 2640  
 DB 2581 AGTACATTCGGAATACCCCTTCTTGAATATTATTATTCGAACATTCCACTGCC 2640  
 QY 2641 CTGGGACCGGTGACCATCTCTGATGACCTTGGGCACTGACATGGTCCCGGCATCTCT 2700  
 DB 2641 CTGGGACCGGTGACCATCTCTGATGACCTTGGGCACTGACATGGTCCCGGCATCTCT 2700  
 QY 2701 CTGGCTATGAAACAGGCTGAAAGTACATCATGAAAGAGCAACCCGAAATCCGAAACG 2760  
 DB 2701 CTGGCTATGAAACAGGCTGAAAGTACATCATGAAAGAGCAACCCGAAATCCGAAACG 2760  
 QY 2761 GACAAACTTTGTAAGAGAGCTGTATCAGACATGAGCTATGACATATGGTATGATCCAG 2820  
 DB 2761 GACAAACTTTGTAAGAGAGCTGTATCAGACATGAGCTATGACATATGGTATGATCCAG 2820  
 QY 2821 GCGCTGGGAGGCTTCTTCACTATTATTTGATTTCTGGCTGAGAACGTTTCTGCTCTT 2880  
 DB 2821 GCGCTGGGAGGCTTCTTCACTATTATTTGATTTCTGGCTGAGAACGTTTCTGCTCTT 2880  
 QY 2881 CACCTGTTGGGACATCCGAGAGACCTGGGATGACCGCTGATTAATGATGTGAGAGACGC 2940  
 DB 2881 CACCTGTTGGGACATCCGAGAGACCTGGGATGACCGCTGATTAATGATGTGAGAGACGC 2940  
 QY 2941 TACGGGACACAGTGAACCTTACGAGAGAGAGAGATTGTGAGATTCACTGACACGCGCC 3000  
 DB 2941 TACGGGACACAGTGAACCTTACGAGAGAGAGAGATTGTGAGATTCACTGACACGCGCC 3000  
 QY 3001 TTCTTTGTCAGTATCCTGCTAGTGTGACATGGGCTGACCTTGGTCACTGCAAGACAGAAAG 3060  
 DB 3001 TTCTTTGTCAGTATCCTGCTAGTGTGACATGGGCTGACCTTGGTCACTGCAAGACAGAAAG 3060  
 QY 3061 AATTCTGTCTTCCAG 3120  
 DB 3061 AATTCTGTCTTCCAG 3120  
 QY 3121 ACAGCTCTTGTCTCTTCT 3180  
 DB 3121 ACAGCTCTTGTCTCTTCT 3180  
 QY 3181 CCGCTTCAAACTTCT 3240  
 DB 3181 CCGCTTCAAACTTCT 3240  
 QY 3241 GACGAGGTGGAGAGCTCATCATCAGGCGAGCGCTGCGGCTGGGTGAGAGAGAGAGAGAG 3300  
 DB 3241 GACGAGGTGGAGAGCTCATCATCAGGCGAGCGCTGCGGCTGGGTGAGAGAGAGAGAGAG 3300  
 QY 3301 TACTACTAGCCCACTGCGCTGCAAGCGCTGCAAGCATTTGTGCTGCACTGCTACCTTACCCC 3360  
 DB 3301 TACTACTAGCCCACTGCGCTGCAAGCGCTGCAAGCATTTGTGCTGCACTGCTACCTTACCCC 3360  
 QY 3361 TACCCCCCTTTGTGTACTTCAAGTCTTGAAGCTCGGAACCTTACCTTGTAGAGAAAGCA 3420  
 DB 3361 TACCCCCCTTTGTGTACTTCAAGTCTTGAAGCTCGGAACCTTACCTTGTAGAGAAAGCA 3420  
 QY 3421 CCAAGCATGTGGGATTCAGACGCTCTGAAATGAAAGCATGTAGCTTAAATGGGGGGCGG 3480  
 DB 3421 CCAAGCATGTGGGATTCAGACGCTCTGAAATGAAAGCATGTAGCTTAAATGGGGGGCGG 3480  
 QY 3481 GGGGAGGGCTGCGCGGAAACACCGTGAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
 DB 3481 GGGGAGGGCTGCGCGGAAACACCGTGAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
 QY 3541 CCTTTTGTGTTTGTAAAG 3600  
 DB 3541 CCTTTTGTGTTTGTAAAG 3600  
 QY 3601 GGATTTTACAAATTAAGATGGCTATTATAAGGAGAA 3636  
 DB 3601 GGATTTTACAAATTAAGATGGCTATTATAAGGAGAA 3636

RESULT 2  
 US-09-090-535-5  
 ; Sequence 5, Application US/09090535A  
 ; Patent No. 6309874  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BELUSA, Roger  
 ; TITLE OF INVENTION: SELECTION MARKER  
 ; FILE REFERENCE: BELUSA 09/090,535  
 ; CURRENT APPLICATION NUMBER: US/09/090,535A  
 ; EARLIER FILING DATE: 1998-06-04  
 ; EARLIER FILING DATE: 1997-06-04  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 3636  
 ; TYPE: DNA  
 ; ORGANISM: Rattus rattus  
 ; US-09-090-535-5

Query Match 99.7%; Score 3626.4; DB 4; Length 3636;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 3630; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAACCTCGGCGGAG 60  
 DB 1 GGAGCTCGGCGGAG 60  
 QY 61 GCGGAG 120  
 DB 61 GCGGAG 120  
 QY 121 CAGGCTTATGTTCCGCGCTCTGAGCTCCCGGCTCCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
 DB 121 CAGGCTTATGTTCCGCGCTCTGAGCTCCCGGCTCCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
 QY 181 CTCTCTTTTCTAGTCTCCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 DB 181 CTCTCTTTTCTAGTCTCCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 QY 241 GGGAGGGGGTTGAG 300  
 DB 241 GGGAGGGGGTTGAG 300  
 QY 301 AAGACAAAGAGGCGAAG 360  
 DB 301 AAGACAAAGAGGCGAAG 360  
 QY 361 GAGACCATTAATCTAGCTGATGAACTCTCAATCGTAAATACGAAACAGATTGAGCCGA 420  
 DB 361 GAGACCATTAATCTAGCTGATGAACTCTCAATCGTAAATACGAAACAGATTGAGCCGA 420  
 QY 421 GGCCTAACACCCGCAAGAGCGCTGAGATCTGAGTCTGAGATGCTGAGATGCTGAGATGCTGAG 480  
 DB 421 GGCCTAACACCCGCAAGAGCGCTGAGATCTGAGTCTGAGATGCTGAGATGCTGAGATGCTGAG 480  
 QY 481 CCGCTCCCACTACTCCGAG 540  
 DB 481 CCGCTCCCACTACTCCGAG 540  
 QY 541 ATGTTACTGTGATGAG 600  
 DB 541 ATGTTACTGTGATGAG 600  
 QY 601 GAAAGAGAAACCAAAATATGATCTGTACCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 DB 601 GAAAGAGAAACCAAAATATGATCTGTACCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 QY 661 ATACTGCGCTGTTCTCTATTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
 DB 661 ATACTGCGCTGTTCTCTATTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
 QY 721 AAGACATGCTCCTCAGCAAGCCCTGTGATTGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 780



```

Db 721 AGAATATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 780
Qy 781 GAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 840
Db 781 GAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 840
Qy 841 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 900
Db 841 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 900
Qy 901 TCAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 960
Db 901 TCAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 960
Qy 961 ATTCCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1020
Db 961 ATTCCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1020
Qy 1021 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1080
Db 1021 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1080
Qy 1081 ACCGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1140
Db 1081 ACCGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1140
Qy 1141 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1200
Db 1141 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1200
Qy 1201 ATTCCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1260
Db 1201 ATTCCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1260
Qy 1261 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1320
Db 1261 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1320
Qy 1321 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1380
Db 1321 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1380
Qy 1381 ACTCAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1440
Db 1381 ACTCAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1440
Qy 1441 ACCGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1500
Db 1441 ACCGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1500
Qy 1501 TCAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1560
Db 1501 TCAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1560
Qy 1561 ATTCCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1620
Db 1561 ATTCCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1620
Qy 1621 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1680
Db 1621 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1680
Qy 1681 TCAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1740
Db 1681 TCAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1740
Qy 1741 AGGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1800
Db 1741 AGGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1800
Qy 1801 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1860
Db 1801 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1860

Db 1801 CTCCTCAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1860
Qy 1861 TACCTAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1920
Db 1861 TACCTAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1920
Qy 1921 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1980
Db 1921 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 1980
Qy 1981 AACCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2040
Db 1981 AACCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2040
Qy 2041 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2100
Db 2041 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2100
Qy 2101 ATTCCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2160
Db 2101 ATTCCTCTCTCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2160
Qy 2161 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2220
Db 2161 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2220
Qy 2221 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2280
Db 2221 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2280
Qy 2281 TACCTAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2340
Db 2281 TACCTAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2340
Qy 2341 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2400
Db 2341 GAGAGAGAGATGTCCTCAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2400
Qy 2401 TACCTAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2460
Db 2401 TACCTAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2460
Qy 2461 TACCTAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2520
Db 2461 TACCTAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2520
Qy 2521 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2580
Db 2521 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2580
Qy 2581 AGTACATTCGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2640
Db 2581 AGTACATTCGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2640
Qy 2641 CTCCTCAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2700
Db 2641 CTCCTCAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2700
Qy 2701 CTCCTCAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2760
Db 2701 CTCCTCAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2760
Qy 2761 GACAGAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2820
Db 2761 GACAGAGAGAGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2820
Qy 2821 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2880
Db 2821 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2880
Qy 2881 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2940
Db 2881 GATCTCAGATATATCTGAAATGAGAGAGCCCTCGTATTCGAAATGAGAGAAAGATGAGATCAAC 2940

```



QY 2941 TACGGGACAGCTGAGCCTACGAGCAGAGAAATTGTGAGATTCACTTGGCAGCAGCGCC 3000  
Db 2941 TACGGGACAGCTGAGCCTACGAGCAGAGAAATTGTGAGATTCACTTGGCAGCAGCGCC 3000  
QY 3001 TTCTTTGTGAGTATCGTGTGATGCGAGTGGGCTGATGTCATCTGCAAGACAGAAAG 3060  
Db 3001 TTCTTTGTGAGTATCGTGTGATGCGAGTGGGCTGATGTCATCTGCAAGACAGAAAG 3060  
QY 3061 AATTGCTGCTTCAGAGAGAAATGAAGAACAGATCTTAATATTGAGCTCTTTGAAGAG 3120  
Db 3061 AATTGCTGCTTCAGAGAGAAATGAAGAACAGATCTTAATATTGAGCTCTTTGAAGAG 3120  
QY 3121 ACAGCTTGTGCTTCTTCT 3180  
Db 3121 ACAGCTTGTGCTTCTTCT 3180  
QY 3181 CCGCTCAAACTACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 3240  
Db 3181 CCGCTCAAACTACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 3240  
QY 3241 GACGAGTGGGAGAGCTCATCATCAGGCGACGCCCTGGCGGCTGGGCTGGAGAGAAACC 3300  
Db 3241 GACGAGTGGGAGAGCTCATCATCAGGCGACGCCCTGGCGGCTGGGCTGGAGAGAAACC 3300  
QY 3301 TACTACTAGGCGCAGTGGCTGGAGCGGCTGGAGCAATTGTGCGACACCTGACCTACCCC 3360  
Db 3301 TACTACTAGGCGCAGTGGCTGGAGCGGCTGGAGCAATTGTGCGACACCTGACCTACCCC 3360  
QY 3361 TACCCCCCTTTGTGACTTCAAGTCTGGAGCTCGGAACTTACCTGTGTAGAAAGCA 3420  
Db 3361 TACCCCCCTTTGTGACTTCAAGTCTGGAGCTCGGAACTTACCTGTGTAGAAAGCA 3420  
QY 3421 CCNAAGCATGTGGGATCCGAGCTCTCTGAAATGAAGCATGTAGCTTATGGGGGCGG 3480  
Db 3421 CCNAAGCATGTGGGATCCGAGCTCTCTGAAATGAAGCATGTAGCTTATGGGGGCGG 3480  
QY 3481 GGGGAGGCTGGCCGAAAAACACCGTGAAGCGGAGCGAGCGGGAAGTTATATGTG 3540  
Db 3481 GGGGAGGCTGGCCGAAAAACACCGTGAAGCGGAGCGAGCGGGAAGTTATATGTG 3540  
QY 3541 CTTTGTGTTTGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3600  
Db 3541 CTTTGTGTTTGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3600  
QY 3601 GGAATTTTAAATTAAGATGGCTATTATTAACGGA 3636  
Db 3601 GGAATTTTAAATTAAGATGGCTATTATTAACGGA 3636

RESULT 3  
US-09-090-535-6  
Sequence 6, Application US/09090535A  
Patent No. 6309874  
GENERAL INFORMATION:  
APPLICANT: BELUSA, Roger  
TITLE OF INVENTION: SELECTION MARKER  
FILE REFERENCE: BELUSA 09/090,535  
CURRENT APPLICATION NUMBER: US/09/090,535A  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: US 60/048,601  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 3636  
TYPE: DNA  
ORGANISM: Rattus rattus  
US-09-090-535-6

Query Match 99.7%; Score 3626.4; DB 4; Length 3636;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3630; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGAAGCTCGGCGGAGAGAGCGGAGACAGTGGCAGCGCGCGGCGGAGCGAGCGAGCG 60  
Db 1 GGAAGCTCGGCGGAGAGAGCGGAGACAGTGGCAGCGCGCGGCGGAGCGAGCGAGCG 60  
QY 61 GCGGACAGAGCGGCGGCTCGGCTCGGAGGCGCGCGGCGGCTCTCTCTCTCTCTCTCT 120  
Db 61 GCGGACAGAGCGGCGGCTCGGCTCGGAGGCGCGCGGCGGCTCTCTCTCTCTCTCTCT 120  
QY 121 CAGCCTAGTTCGCGCTCTCGGCTCTCGGCTCTCGGCTCTCGGCTCTCGGCTCTCGGCT 180  
Db 121 CAGCCTAGTTCGCGCTCTCGGCTCTCGGCTCTCGGCTCTCGGCTCTCGGCTCTCGGCT 180  
QY 181 CTGCTCTTCTAGTCTCCAGCCAGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
Db 181 CTGCTCTTCTAGTCTCCAGCCAGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
QY 241 GGGAGGGGCTTGAAGAGAGCAAGTATGAGCGCGCAGCTGTATCAGAAATGAGAGAG 300  
Db 241 GGGAGGGGCTTGAAGAGAGCAAGTATGAGCGCGCAGCTGTATCAGAAATGAGAGAG 300  
QY 301 AAGAGCAAGAGCGGAG 360  
Db 301 AAGAGCAAGAGCGGAG 360  
QY 361 GAGGACCATTAATCTCAGCTGATGAACTCCATCTTAATACGAAAGAGAGAGAGAGAG 420  
Db 361 GAGGACCATTAATCTCAGCTGATGAACTCCATCTTAATACGAAAGAGAGAGAGAGAG 420  
QY 421 GGGCTAACACCCGCGAGGCGGCTGAGATCTCTGCTGGGATGGCGGCGGCGGCGGCT 480  
Db 421 GGGCTAACACCCGCGAGGCGGCTGAGATCTCTGCTGGGATGGCGGCGGCGGCGGCT 480  
QY 481 CCGCTCCGCTACTCCGAGTGGGTCAAAATCTGTGGGAGCTGTGTGGGTGGGTGGGT 540  
Db 481 CCGCTCCGCTACTCCGAGTGGGTCAAAATCTGTGGGAGCTGTGTGGGTGGGTGGGT 540  
QY 541 ATGTACTGTGATGGAGGCTATCTTTGTTCTTGCTTATGAGCATTCGAGTGTCTACA 600  
Db 541 ATGTACTGTGATGGAGGCTATCTTTGTTCTTGCTTATGAGCATTCGAGTGTCTACA 600  
QY 601 GAAGAGACCAACCAATGATGATCTGATCTGGGAGTGGGTGCTGCTGCTGCTGCTGCT 660  
Db 601 GAAGAGACCAACCAATGATGATCTGATCTGGGAGTGGGTGCTGCTGCTGCTGCTGCT 660  
QY 661 ATTAAGTGGTGTCTCTCTATTAATCAAGAGCAAAAGCTCAAGATCATGAAATCTCTTC 720  
Db 661 ATTAAGTGGTGTCTCTCTATTAATCAAGAGCAAAAGCTCAAGATCATGAAATCTCTTC 720  
QY 721 AAGAACTGTCTCTCTCAAGAGGCTCTGATTTGAAATGGAAGAGATGAGATGAGATCAAC 780  
Db 721 AAGAACTGTCTCTCTCAAGAGGCTCTGATTTGAAATGGAAGAGATGAGATGAGATCAAC 780  
QY 781 GCAGAGATGCTGCTGTGATCTGGTGAAGTGAAGGCGGAGACCGAATCCCTGCT 840  
Db 781 GCAGAGATGCTGCTGTGATCTGGTGAAGTGAAGGCGGAGACCGAATCCCTGCT 840  
QY 841 GATCTCAGATCATATGTGCAATGCTGCAAGGTGATTAATCTCTCACTACTGCTGATA 900  
Db 841 GATCTCAGATCATATGTGCAATGCTGCAAGGTGATTAATCTCTCACTACTGCTGATA 900  
QY 901 TCAGAACCCGAGACTCGGCTCGGCTCGGATTTCAAAAGAGAGAGAGAGAGAGAGAGAG 960  
Db 901 TCAGAACCCGAGACTCGGCTCGGCTCGGATTTCAAAAGAGAGAGAGAGAGAGAGAGAG 960  
QY 961 ATTGCTTCTTCTCAACCAACTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
Db 961 ATTGCTTCTTCTCAACCAACTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
QY 1021 GGGGATCGCAGCTGATGGAG 1080  
Db 1021 GGGGATCGCAGCTGATGGAG 1080







Db 3241 GACGAGGTGCGGAGAGCTATCATCAGCGAGCGCCTGCGGCTGGGTGGAGGAAGAACCC 3300  
QY 3301 TACTACTAGCCCACTGCGCCCTGCAGCGCGTGGAAATTGCGCACACACTGCACTATACCC 3360  
Db 3301 TACTACTAGCCCACTGCGCCCTGCAGCGCGTGGAAATTGCGCACACACTGCACTATACCC 3360  
QY 3361 TACCCCTCTTTGTGTACTTCAAGTCTTGGAGCTCGGAACCTGACCTGTTAGGAAGA 3420  
Db 3361 TACCCCTCTTTGTGTACTTCAAGTCTTGGAGCTCGGAACCTGACCTGTTAGGAAGA 3420  
QY 3421 CCAAGCATGTGGAGATCTTAAACCTGCTGGAATGAAGCATGTAGCTGTAAATGCGG 3480  
Db 3421 CCAAGCATGTGGAGATCTTAAACCTGCTGGAATGAAGCATGTAGCTGTAAATGCGG 3480  
QY 3481 GGGAGGAGGTGCGCCCAAGAACCGTGGAGCGGAGAGCAAGCGGGAGAGTTATATGTG 3540  
Db 3481 GGGAGGAGGTGCGCCCAAGAACCGTGGAGCGGAGAGCAAGCGGGAGAGTTATATGTG 3540  
QY 3541 CCTTTTGTGTGTAAAGGAAAACTGGAAAGACTGAAGATTAAGTTATATATCT 3600  
Db 3541 CCTTTTGTGTGTAAAGGAAAACTGGAAAGACTGAAGATTAAGTTATATATCT 3600  
QY 3601 GGATTTTAAAGATAAGTGGCTATTATACGGAA 3636  
Db 3601 GGATTTTAAAGATAAGTGGCTATTATACGGAA 3636

## RESULT 4

US-09-090-535-7  
Sequence 7, Application US/09090535A  
Patent No. 6309874  
GENERAL INFORMATION:  
APPLICANT: BEIUSA, Roger  
TITLE OF INVENTION: SELECTION MARKER  
FILE REFERENCE: BEIUSA 09/090,535  
CURRENT FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: US 60/048,601  
EARLIER FILING DATE: 1997-06-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 3636  
TYPE: DNA  
ORGANISM: Rattus rattus  
US-09-090-535-7

Query Match 99.6%; Score 3621.6; DB 4; Length 3636;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3627; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GGAGCCTCGGCGGAGAGAGCGGAGCACTGGCAGCGGCGGCGGAGCGAGCAGCGG 60  
Db 1 GGAGCCTCGGCGGAGAGAGCGGAGCACTGGCAGCGGCGGCGGAGCGAGCAGCGG 60  
QY 61 GCGGAGCAGCGCGGCGCTCGGTCCGCGGCGCGCGGCTCTCTCTCTCTCTCTCTCT 120  
Db 61 GCGGAGCAGCGCGGCGCTCGGTCCGCGGCGCGCGGCTCTCTCTCTCTCTCTCTCT 120  
QY 121 CAGCCCTAGTCTCCGCTCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
Db 121 CAGCCCTAGTCTCCGCTCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
QY 181 CT 240  
Db 181 CT 240  
QY 241 GCGAAGGCGGTTGAGCAGACAGATATGAGCCCGCAGCTGTATCAAGCAATGCGGCAAG 300  
Db 241 GCGAAGGCGGTTGAGCAGACAGATATGAGCCCGCAGCTGTATCAAGCAATGCGGCAAG 300  
QY 301 AAGAGCAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
Db 301 AAGAGCAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360

Db 301 AAGAGCAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
QY 361 GAGGACATAAATCTAGCCTGATGACTCCATCGTAATATGGAACAGATTGAGCGCA 420  
Db 361 GAGGACATAAATCTAGCCTGATGACTCCATCGTAATATGGAACAGATTGAGCGCA 420  
QY 421 GGCCTAACCCGCGAGAGGCGCTGAGATCTGTCGCGGATGAGCCCAAGCGCTTCAGG 480  
Db 421 GGCCTAACCCGCGAGAGGCGCTGAGATCTGTCGCGGATGAGCCCAAGCGCTTCAGG 480  
QY 481 CCGCTCTCCACTACTCTCCGAGTGGTCAAAATCTGTGCGGCACTGTTCGCTCTCTCC 540  
Db 481 CCGCTCTCCACTACTCTCCGAGTGGTCAAAATCTGTGCGGCACTGTTCGCTCTCTCC 540  
QY 541 ATGTTACTGTGATGAGCAATCTCTGTTCTTGGCTTATGAGTCCGAATGCTGAC 600  
Db 541 ATGTTACTGTGATGAGCAATCTCTGTTCTTGGCTTATGAGTCCGAATGCTGAC 600  
QY 601 GAAGAGAACCAACCAATGATGATCTGTACCTCGGGTCTGCTGTCTGCTCATC 660  
Db 601 GAAGAGAACCAACCAATGATGATCTGTACCTCGGGTCTGCTGTCTGCTCATC 660  
QY 661 ATATCTGCTGTTTCTCTCTATTATCAAGAGCAAAAGCTCCCAAGATCATGAACTCTTC 720  
Db 661 ATATCTGCTGTTTCTCTCTATTATCAAGAGCAAAAGCTCCCAAGATCATGAACTCTTC 720  
QY 721 AAGACATGATGCTCTGAGAGGCGCTGATTCGAATGAGAGAGATGACATCAAC 780  
Db 721 AAGACATGATGCTCTGAGAGGCGCTGATTCGAATGAGAGAGATGACATCAAC 780  
QY 781 GCAGAGATGTCGTCTGTGTGTATCTGTGTGAGAGTGAAGGCGGAGACCGAATCCCTGCT 840  
Db 781 GCAGAGATGTCGTCTGTGTGTATCTGTGTGAGAGTGAAGGCGGAGACCGAATCCCTGCT 840  
QY 841 GATCTGAGATCATATCTGCAAAATGCTGCAAGGTGATTAATCTCTCTCTCTCTCTCTCT 900  
Db 841 GATCTGAGATCATATCTGCAAAATGCTGCAAGGTGATTAATCTCTCTCTCTCTCTCTCT 900  
QY 901 TCAGAACCCGAGCTCGGTCGCCGAGTTTCACAAAGAGAACCCCTTGGAGACAAAGAAC 960  
Db 901 TCAGAACCCGAGCTCGGTCGCCGAGTTTCACAAAGAGAACCCCTTGGAGACAAAGAAC 960  
QY 961 ATTGCTTCTTCTCAACCAACTGTGTGAAAGAACTGCACTGCTGTGTGTACT 1020  
Db 961 ATTGCTTCTTCTCAACCAACTGTGTGAAAGAACTGCACTGCTGTGTGTACT 1020  
QY 1021 GGGGATCGACCGGTGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 1080  
Db 1021 GGGGATCGACCGGTGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 1080  
QY 1081 ACCCCATTGCTGAAGAAATCGAGCACTTCATCAGGCTGATCAGGCTGCTGCTTC 1140  
Db 1081 ACCCCATTGCTGAAGAAATCGAGCACTTCATCAGGCTGATCAGGCTGCTGCTTC 1140  
QY 1141 CTGGGGGTCTCTTCTTCAATCTCTCTCTGATCTTGAAGTCACTTGTGAGCTGTCT 1200  
Db 1141 CTGGGGGTCTCTTCTTCAATCTCTCTCTGATCTTGAAGTCACTTGTGAGCTGTCT 1200  
QY 1201 ATCTTCCTATGATGATCATCTGATGAGCAACGTCGCGGAGGCTTGTCTGCGCAACGTCAG 1260  
Db 1201 ATCTTCCTATGATGATCATCTGATGAGCAACGTCGCGGAGGCTTGTCTGCGCAACGTCAG 1260  
QY 1261 GATATGCTGACGCTCACTGCAAGCGCATGCGAGGAAGAACTGCTGTGAGAACTCTG 1320  
Db 1261 GATATGCTGACGCTCACTGCAAGCGCATGCGAGGAAGAACTGCTGTGAGAACTCTG 1320  
QY 1321 GAACTGTGAGACCTTGGGCTCACTCCACATCTGCTCGCAAGACTGGAACCTCTG 1380  
Db 1321 GAACTGTGAGACCTTGGGCTCACTCCACATCTGCTCGCAAGACTGGAACCTCTG 1380  
QY 1381 ACTCAGAACCGAGTGAAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
Db 1381 ACTCAGAACCGAGTGAAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440







Db 3601 GGATTTTACAATAAGATGGCTATTATAACGGA 3636

## RESULT 5

```

1 Sequence 501: Application US/09385982
2
3 Patent No. 6262394
4
5 GENERAL INFORMATION:
6
7 APPLICANT: ENDEGE, WILSON O., ET AL.
8
9 TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
10
11 TITLE OF INVENTION: PRODUCTS: II
12
13 FILE REFERENCE: CCNDA-260XX
14
15 CURRENT APPLICATION NUMBER: US/09/385,982
16
17 CURRENT FILING DATE: 1999-08-30
18
19 EARLIER APPLICATION NUMBER: 09/328,111
20
21 EARLIER FILING DATE: 1999-06-08
22
23 EARLIER APPLICATION NUMBER: 60/117,393
24
25 EARLIER FILING DATE: 1999-01-27
26
27 EARLIER APPLICATION NUMBER: 60/098,639
28
29 EARLIER FILING DATE: 1998-08-31
30
31 NUMBER OF SEQ ID NOS: 544
32
33 SOFTWARE: fastSeq for Windows Version 3.0
34
35 SEQ ID NO 501
36
37 LENGTH: 605
38
39 TYPE: DNA
40
41 ORGANISM: Homo sapiens
42
43 FEATURE:
44
45 NAME/KEY: misc_feature
46
47 LOCATION: (1)_(605)
48
49 OTHER INFORMATION: n = A,T,C or G
50
51 US-09-385-982-501

```

Query Match	9.5%;	Score 344.4;	DB 4;	Length 605;
Best Local Similarity	85.3%;	Pred. No. 6.7e-81;		
Matches 428;	Conservative 0;	Mismatches 66;	Indels 8;	Gaps 4

QY	2288	ACCAACGGGAATATGTTCTTTGGCTAGAGACCTCTCCTCAACAGAAAGCTCATATTGGAGAG	2347
Db	1	ACCAACTCTGAGATATGTTGTTTCCAGAACCTCCCTCAAGCAGAAAGCTCATATTGGAGAG	60
QY	2348	GCTGTCAGGCGGACAGGAGTGCATATGTGTGCTGTCAACAGGGAGTGTGTCAATGACTCTCCAG	2407
Db	61	GCTGTCAGGCGGACAGGAGTGCATATGTGTGCTGTCACTGTGTACCGTGTGAATGACTCTCCAG	120
QY	2408	CTTTGAAAAAAGCAGATATTGGGGGTTGCCATGGGGAGTTGTTGGCTCGATGTGTCCAAAG	2467
Db	121	CTTTGAAAGAAAGCAAATATTGGGATTCATAGGSAATTGCTGGCTCAGATATGTCCAAAG	180
QY	2468	AAGCTGCTGACATGATTTCTTCTGGATGACAACTTTGCTTCATCGTACTGGAGTAGAAG	2527
Db	181	AAGCTGCTGACATGATTTCTTCTGGATGACAACTTTGCTTCATTTGACTGGAGTAGAAG	240
QY	2528	AAGTTCGTCTGATATTGTAATGATCTGTAAGAAATCCAAATGTTACACCCCAACAGATACA	2587
Db	241	AAGTTCGTCTGATATTGTAATGATCTGTAAGAAATCCAAATGTTATACCTTAACAGATACA	300
QY	2588	TT-CCGSAATCACCCCTTCTTGAATATTATATTGCAAACTTCCACTGCCCTCGGC	2646
Db	301	TTCCCGCGAGATCACCCGTTCTCTGATTTATTATGCAAACTTCCACTCACACTCGGG	360
QY	2647	ACCGTGACCATCTCTGTGATTTGACTTGGGACATGACATGATTTGCCGCATCTCTGCGC	2706
Db	361	ACTGTACCATCTCTGTGATTTGACTTGGGACATGACATGAGGATCTGCATCT-CTTGGCT	419
QY	2707	TATGAACAGGCTGAAAGTGAATGATGATGAAAGAGCGACGCCAGAAATCCCAAAACGACAAA	2766
Db	420	TATGAGCAGGCTGGA--GGGCATCATGAANAAACAGGCCCAAAATCC-----AAACAGACAA	473
QY	2767	CTTGTGAACGAGCGCTTGATTA	2788
Db	474	CTTGTGAATGAACGCGTGTGATTA	495

RESULT 6  
US-08-96

```

US-08-961-527-55/C
Sequence 55, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 21040 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```

Query Match	3.7%;	Score 133;	DB 4;	Length 21040;
Best Local Similarity	53.9%;	Pred. No. 1.5e-24;		
Matches 321;	Conservative	0;	Mismatches 265;	Indels 9;
				Gaps 2

QY	2142	AAGCATTGGCCGCCCGCTCCATTCGCAATGAAACGATGTAACCTCCAGAGATGCGCAAG	2231
Db	1875	AAGACACAGCGGAAGCCATTGCCAAACGTTTGGGAATCATTTACGCCAACAAGATACAGAAG	1816
QY	2222	CCCTGTAGTACATGCGAGTGACTTGAAGGACATGACCTCTTGAGAGCTGGATGACATTT	2281
Db	1815	GTCACGTTTAACTGGGTGGAACCTACATGAACGTGCAGATGAAGATTTGAAAAAGTCG	1756
QY	2282	TGCGGTACCAACGGGAATTTGCTTTGGTAGAACCTCTCTCTAAACGAAGGCTCATCATTTG	2341
Db	1755	TTGGCAATACTC-----TGTATTAGCCCTGTGTCTCTCAAGAACACAAAGTTGGTATCG	1702
QY	2342	TGGAGGGCTGGCCACGCGCAGGCTGTGCATCTGTGTGTGTACAGGAGATGAGTGTCAATGACT	2401
Db	1701	TGCAAGCTTGGCAAAAAACAAGTAAAGCTGTTCGCATGACAGGTGACGGTGCATGTACAG	1642
QY	2402	CTCCAGCTTTAAAAAAGACGATATTTGGGTTTCCCATGGGGAATTTGTTCGTGTGATTTGTT	2461
Db	1641	GCCACGCTCTAAAAACAGCCGATATCGTATCGTATGGGAATCACTGTGTACAGAGGTTT	1582
QY	2462	CCAAACAGAGCTGTGACATGATCTTTCTGTGANTACAACTTTGGCTTCATCGTGACTGGAG	2521
Db	1581	CTAAGGGGGCTTCGATATGATTTCTTGACATGATTAACCTTGGACCTATTATTCGTGGAG	1522
QY	2522	TGAGAGAGGTGCTGTGATATTTGATTAACCTTGAAGAAATTCATTTGTTACACCTTACAA	2581







QY 2292 CACCGAGATTGTCCTTTGCTAGGACCTTCTCTCAACAGAGCTCATCATTTGTGGAGGCTG 2351







Downloaded from <http://ajph.org/> on November 10, 2014

100

100







GenCore version 5.1.4 ps\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2003, 19:45.36, Search time 233 Seconds  
(without alignments)  
10422.991 Million cell updates/sec

Title: US-10-040-722-3

Perfect score: 3636

Sequence: 1 ggaagctcgagcgagagag... agatgcataataacgaa 3636

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 478924 seqs, 33395956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US05\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US04\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US03\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US02\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US01\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US00\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US12\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US13\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1320.4	36.3	1338	10	US-09-870-472-7
2	1127.4	31.0	3053	10	US-09-834-975-967
3	870.8	23.9	1338	10	US-09-870-472-9
4	844.4	23.2	1338	10	US-09-870-472-8
5	780.2	21.5	1355	10	US-09-870-472-11
6	779.8	21.4	1355	10	US-09-870-472-12
7	750.4	20.6	1355	10	US-09-870-472-10
8	486	13.4	569	9	US-10-046-935-473
9	486	13.4	569	9	US-09-878-178-809
10	432.6	11.9	533	9	US-10-076-622-271
11	432.6	11.9	533	10	US-09-604-287A-271
12	432.6	11.9	533	10	US-09-339-338-271
13	432.6	11.9	533	12	US-10-007-805-271
14	367	10.1	447	9	US-09-796-692-8244
15	343.2	9.4	420	10	US-09-960-352-6966
16	330	9.1	432	10	US-09-360-352-6292
17	293.4	8.2	332	10	US-09-960-352-1493
18	290.4	8.0	364	9	US-10-046-935-890
19	290.4	8.0	364	9	US-09-878-178-890

20	267	7.3	415	10	US-09-960-352-6766	Sequence 6766, App
21	234.6	6.5	468	10	US-09-864-761-6168	Sequence 6168, App
22	228.6	6.3	768	10	US-09-910-943-33	Sequence 33, App1
23	225.8	6.2	269	10	US-09-864-761-21979	Sequence 21979, A
24	209.6	5.8	3053	10	US-09-834-975-967	Sequence 967, App
25	171.8	4.7	26668	10	US-09-962-635-222	Sequence 222, App
26	164	4.5	215	10	US-09-834-975-485	Sequence 485, App
27	151.2	4.2	487	10	US-09-864-761-5353	Sequence 5353, App
28	149.4	4.1	797	10	US-09-834-975-299	Sequence 299, App
29	147	4.0	150	10	US-09-864-761-22127	Sequence 22127, A
30	142	3.9	2488	10	US-09-974-300-2016	Sequence 2016, App
31	133.6	3.7	2676	9	US-09-738-626-1708	Sequence 1708, App
32	127.2	3.5	13932	10	US-09-070-927A-363	Sequence 363, App
33	121.4	3.3	14286	10	US-09-070-927A-162	Sequence 162, App
34	117.4	3.2	1493	9	US-09-925-299-180	Sequence 180, App
35	117.4	3.2	1493	10	US-09-925-299-180	Sequence 180, App
36	117.2	3.2	534	10	US-09-974-300-6433	Sequence 6433, App
37	117.2	3.2	831	10	US-09-822-830A-406	Sequence 406, App
38	110.4	3.0	1939	10	US-09-822-830A-406	Sequence 1957, App
39	105.2	2.9	3186	9	US-09-938-842A-1957	Sequence 1957, App
40	103.8	2.9	460	10	US-09-960-352-10862	Sequence 10862, A
41	103.2	2.8	3210	9	US-09-938-842A-1594	Sequence 1594, App
42	102.8	2.8	895	10	US-09-070-927A-374	Sequence 374, App
43	94.8	2.6	3063	9	US-09-938-842A-2553	Sequence 2553, App
44	94	2.6	181	10	US-09-833-381-1505	Sequence 1505, App
45	86.2	2.4	968	10	US-09-833-381-641	Sequence 641, App

## ALIGNMENTS

RESULT 1  
US-09-870-472-7  
Sequence 7, Application US/09870472  
Patent No. US20020095023A1  
GENERAL INFORMATION:  
APPLICANT: LEE, Kyunglim et al.  
TITLE OF INVENTION: Ige-dependent histamine-releasing factor (HRF) receptor, HRF-  
TITLE OF INVENTION: Binding peptides, nucleic acids encoding the same, and uses  
FILE REFERENCE: 1599-0198P  
CURRENT APPLICATION NUMBER: US/09/870,472  
CURRENT FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: KR10-2000-0030130  
PRIOR FILING DATE: 2000-06-01  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: KOPATIN 1.71  
SEQ ID NO 7  
LENGTH: 1338  
TYPE: DNA  
ORGANISM: Rattus sp.  
US-09-870-472-7

Query Match 36.3%; Score 1320.4; DB 10; Length 1338;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1327; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1222 GTAGCCCAAGCGCGGAAAGTTTCTGGCCACCGTCAAGTATGTCAGCGCTCACTGCC 1281  
1 GTAGCCCAAGCGCGGAAAGTTTCTGGCCACCGTCAAGTATGTCAGCGCTCACTGCC 60  
1:82 AAGGCATGGCGAGAAAGAACTGCTGTGAAGAACTGGAAGCTGGAGACCTTGGGG 1341  
61 AAGGCATGGCGAGAAAGAACTGCTGTGAAGAACTGGAAGCTGGAGACCTTGGGG 120  
1342 TCCGATCCACATCTGCTCCGACAGACTGGAATCTGATCGAACCAGATGACAGTG 1401  
121 TCCGATCCACATCTGCTCCGACAGACTGGAATCTGATCGAACCAGATGACAGTG 180  
1402 GCTACATGTTGTTGACATCAATTCATGAGTGAACACCAAGAGAAATCAGAGTGGG 1461  
181 GCTACATGTTGTTGACATCAATTCATGAGTGAACACCAAGAGAAATCAGAGTGGG 240







QY 2682 TGCATTGACTTGGGACCTGACATGTTTCCCGCAGCTCTCGGGCTCTFAGAAACAGGGCTGAA 2721  
 Db 619 TGCATTGACTTGGGACCTGACATGTTTCCCGCAGCTCTCGGGCTCTFAGAAACAGGGCTGAG 678  
 QY 2722 AGTGACATCATGAAACAGGAGGCCCGAAGATCCCAAAACGGAACAACCTTGTAACGAGCGT 2781  
 Db 679 AGTGACATCATGAAACAGGAGGCCCGAAGATCCCAAAACGGAACAACCTTGTAATGAGCGG 738  
 QY 2782 CTGATCAGCATGCGCTATGACAGATCGGTATGATCCAGGGCCCTGGGAGGCTTCTTACT 2841  
 Db 739 CTGATCAGCATGCGCTATGAGCAGATTGGAAATGATCCAGGGCCCTGGGAGGCTTCTTACT 798  
 QY 2842 TATTTTGTCATTCTGGCTGAGAACGGTTTCCGCCCTTTCACCTGTTGGGATCCGAGAG 2901  
 Db 799 TACTTTGTCATTCTGGCTGAGAACGGCTTCTCCCAATTACACTGTTGGGCTCCGAGTG 858  
 QY 2902 ACCTGCGATGACCCGTGTGAAATTATATGTTGAGAGACAGTACGGGACAGCTGAGACTAC 2961  
 Db 859 GACTGGGATGACCGCTGGATTCACACATGTGGAAGCAGTCTGACACATGTGACCTAT 918  
 QY 2962 GACCAAGAGAAATTTGTGAGATTCACTTGCACACAGGCTTCTTGTGATGATGTTGTA 3021  
 Db 919 GACCAAGAGAAATTTGTGAGATTCACTTGCACACAGCTTCTTGTGATGATGTTGTA 978  
 QY 3022 GTGCAGTGGGCTGACTTGTGATCTGTGCAAGACCAAGAGAAATTCTGCTTCCAGCAGGGA 3081  
 Db 979 GTGCAGTGGGCTGACTTGTGATCTGTGCAAGACCAAGAGAAATTCTGCTTCCAGCAGGGA 1038  
 QY 3082 ATGAGAGAACAGATCTTAATATTTGGCCCTCTTGAAAGACAGCTCTTGCGCTTCCG 3141  
 Db 1039 ATGAGAGAACAGATCTTAATATTTGGCCCTCTTGAAAGACAGCTCTTGCGCTTCCG 1098  
 QY 3142 TCTTACTGCGCTGGATGATGATGTGAGCCCTTAGAGTATCCCTCAAACTACTGTGG 3201  
 Db 1099 T----- 1099  
 QY 3202 TTCTGTGCTTCCCTTACTCCCTTCTCATCTTGTGTATGACAGAGTGGGAACTCATC 3261  
 Db 1100 -----CCTACTCTCTTCTCATCTTGTGTATGACAGAGTGGGAAACTCATC 1148  
 QY 3262 ATCAGCGACAGCGCTGCGGCTGGGTGAGAAAGAACTTACTACTAGCCGCACTGCTG 3321  
 Db 1147 ATCAGCGACAGCGCTGCGGCTGGGTGAGAAAGAACTTACTACTAGCCGCACTGCTG 1206  
 QY 3322 CACGCGGTGAAACATTGTGCCACACT--GCACCTACCCCTTACCCCTTTGTGTGAC 3378  
 Db 1207 CACGCGGTGAAACATTGTGCCACACTGTGCATCCGACACCCACCCCTCTTGTGTGAC 1266  
 QY 3379 TTCAAGCTTGGAGCTGCGGAACCTACCTCGTATGAGAAAGACCAAGAGATGTGGG--AT 3433  
 Db 1267 TTCAAGCTTGGAGCTTGGAGACCTTACCTCGTATGAGAAAGACCAAGAGATGTGGGAA 1322  
 QY 3438 CCAGACGTCCTGGAATGAGAGATGTAGCTGTAATGGGGGGGGGGGAGAGGCTGCCGGA 3497  
 Db 1326 CAGACGTCCTGGAATGAGAGATGTAGCTGTAATGGGGGGGGGGGAGAGGCTGCCGGA 1385  
 QY 3498 AAACACCGTGAACGGGAGACGACAGCGGGGAAGGTTTATAGTCCCTTTTGTGTTGTA 3557  
 Db 1386 ACCATTCATCTGTGGAATGACACCGGGGAAGGTTTATAGTCCCTTTTGTGTTGTA 1444  
 QY 3558 AAAAGAGAAACCTGGAAAGATGTAAGATTACGTTTATATCTGATATTTTACAAATAA 3617  
 Db 1446 AAAAGAGAAACCTGGAAAGATGTAAGATTACGTTTATATCTGATATTTTACAAATAA 1506  
 QY 3618 GATGGCTATTATTAACGAA 3636  
 Db 1506 GATGGCTATTATTAAATGAA 1524

RESULT 3  
 US-09-870-472-9  
 Sequence 9, Application US/09870472  
 Patent No. US20090950381

```

GENERAL INFORMATION:
APPLICANT: LEE, Kyunglim et al.
TITLE OF INVENTION: Igb-dependent histamine-releasing factor(HRF) receptor, HRF-
TITLE OF INVENTION: binding peptides, nucleic acids encoding the same, and uses
TITLE OF INVENTION: thereof
FILE REFERENCE: 1599-0198P
CURRENT APPLICATION NUMBER: US/09/870,472
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: KR10-2000-0030130
PRIOR FILING DATE: 2000-06-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: KOPATIN 1.71
SEQ ID NO 9
LENGTH: 1338
TYPE: DNA
ORGANISM: Rattus sp.
US-09-870-472-9

Query Match      23.9%;   Score 970 3;   DB 10;   Length 1338;
Best Local Similarity 78.2%;   Pred. No. 3,5e+217;
Matches 1046;   Conservative 0;   Mismatches 292;   Indels 0;   Gaps 0;

```

Query Match	23.9%	Score 470.8	DB 10	Length 138	
Best Local Similarity	78.2%	Pred. No. 3.5e-217			
Matches 1046	Conservative	0	Mismatch 252	Indels 0	Gaps
Qy 1222	GTAGCCAACTGCGGAAAGATTGGCTGGGCACCGTCACGGATATGTCTGACGGCTCACTGCC	1281			
Db 1	GTGGCCAAATCTCCCAAGGGGGCGTGGCTTACTGTGCAGGGTGTGCTGACGCTGACGCC	60			
Qy 1282	AAGCGCATGGCGGAGAAAGACTGCTCTGTGAAGAACTTGGAAAGCTGTGGAAACCTTGGGG	1341			
Db 61	AAGCGCATGGCTCGGAAAGAACTGTCTGGTAAAGAACTGGAGGCGGTGGAGAACTTGGGC	120			
Qy 1342	TTCGACATCCCAACATCGCTGCCGACAAAGACTGGAACTGTGACTCAGAACCCGATACAGTG	1401			
Db 121	TTCACATCCACCACTGGCTCTCCAGAAAGCCGCACTCCACCCGAAACCGCATACCGTC	180			
Qy 1402	GCTCACATGTGTGTGTTGACATCAATCAATTCATGAAGCTGACACCAAGAGATACAGTGGG	1461			
Db 181	GCCCACTGTGTGTGTTGACCAACGAGATCCACAGGCGCGACACTACAGAGATACAGTCAAGG	240			
Qy 1462	GTCCTCTTTGACAAAGACGTAGAGCACTACCTGGTTCGCTGTGTCAGAAATTTGCTGTCTGT	1521			
Db 241	ACCTCTTTGACAAAGACTACACACTGGGTGGCCGTGTCCACATGGCGGTCTTGGC	300			
Qy 1522	AACAGGAGATGTTTCAGGCTTAAACCAAGAAAACTTGCCATTCCTTAAAGCGTGGCAAGG	1581			
Db 301	AACCGGCGTGTCTTCAAGGGCGGGGACAGATATACATCCCTGTACTCAAGAGGAGACTGTGGG	360			
Qy 1582	GGAATATCTTCCGATCGCGGCGCTCTTAAAGTGATAGAGGTCTGTGTGGCTCCCGTATG	1641			
Db 361	GATATGCTCTAGAGTCCGGCTGCTTAAGTGGATGAGACTGTCTCGGGTTCCTTAAG	420			
Qy 1642	GAGATGAGGGGGAAGTACACCAAGATATAGTGAATTCCTTCACTCCACCAACAGATAC	1701			
Db 421	CTGATGCCCGAAGCAAAACAAAGAAAGTGGCCGAGATTCCTTCACTCCACATCAAAATAC	480			
Qy 1702	CAGCTTCCTATTCACAAAGAACCCAAACGATCGGAGCTTAAGCACTGTAGTATGAAG	1761			
Db 481	CAGCTATCCATCCATGAGACTGAGAGACCCCAATGACAACCGATACCTGTTAGTATGATGAAG	540			
Qy 1762	GCGGCCCAAGAAAGATCTTGAACCAATGACATTTATCTCTCTCCAGGGCAAGAGAGAG	1821			
Db 541	GCGGCCCTCGAAGCATTTCTTGGACCGCTGGGACCATCTCTCTGAGAGGCAAGAGAGAG	600			
Qy 1822	CGCTTGAGCAAGAGAGTGAAGACGCTTTCAGATATGCTTAAAGCTAGAGCTGGAGGGCTT	1881			
Db 601	CGCTTGAGTGAAGATGAAAGAGAGCCTTCCAGAAAGCCCTACTGAGAGCCCTGAGTGGGCTG	660			
Qy 1882	GGAGAGCGTGTGCTAGGTTTCTGCCACCTCTTCTGTGCTGACGAAAGATTTCCCGAAGGC	1941			
Db 661	GGCGAGCGTGTGCTGGTTTCTGCCATTACTACTGCGGAGGAAAGATTTCCCGAAGGC	720			
Qy 1942	TTCGATTTGACATGATGAAGTCAATTTCCCGGTGATTAACCTGTGCTGTGGTCTTT	2001			
Db 721	TTTGCTTTGATGTGATGACGTAACTTCAACCAAGAACACTTTGCTTGGTGGTCTTC	780			







RESULT 6  
US-09-870-472-12  
; Sequence 12, Application US/09870472  
; Patent No. US20020095023A1















NUMBER OF SEQ ID NOS: 489  
SOFTWARE: fastseq for Windows Version 3.0  
SEQ ID NO: 271  
LENGTH: 533  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1) (533)  
OTHER INFORMATION: n=A,T,C or G  
US-09-604-287A-271

Query Match 11.9%; Score 432.6; DB 10; Length 533;  
Best Local Similarity 88.0%; Pred. No. 5.3e-103;  
Matches 468; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1250 CCACCGTCACGATATGTCAGCTCACTGCCAAGCGGATGCGCAGGAAAGAACTGCTGG 1309  
DB 1 CCACTGTACGGTCTGTCTGACACTTACTGCCAAAAGCATGCGCAAGAAAAAAGCTGTAG 60  
QY 1310 TGAAGAACCTGGAAGCTGTGGAGACCTTGGGGTCCACATCCACCATCTGCTCCGACAGA 1369  
DB 61 TGAAGAACCTTGAAGAGCTGTGGAGACCTTGGGGTCCACATCCATCTGCTGTATAAA 120  
QY 1370 CTGGAACCTGACTCAGAACCGGATGACAGTGGCTCAGATGGTTGACAAATCAATCC 1429  
DB 121 CTGGAACCTGACTCAGAACCGGATGACAGTGGCTCAGATGGTTGACAAATCAATCC 180  
QY 1430 ATGAAGCTGACACCAAGAGAAATGAGAGTGGAGTCTCTTTGACAAAGCTGACCACT 1489  
DB 181 ATGAAGCTGATACACAGAGAAATGAGAGTGGTGTCTTTTGAACAAGCTTCACTACT 240  
QY 1490 GGTTCGCTCTGTCAGAAATGCTGCTCTCTTAACAGGAGAGTGTTCAGGCTAACCAAG 1549  
DB 241 GGTTCGCTCTGTCAGAAATGCTGCTCTCTTAACAGGAGAGTGTTCAGGCTAACCAAG 300  
QY 1550 AAAACCTGCTATCTTAAAGCTGACAGTGGAGAGATGCTTCCAGTGGGCTCTTAA 1609  
DB 301 AAAACCTGCTATCTTAAAGCTGACAGTGGAGAGATGCTTCCAGTGGGCTCTTAA 360  
QY 1610 AGTCATGAGAGT 1669  
DB 361 AGTCATGAGAGT 420  
QY 1670 TGGAGATTCCTTCACTCCACCAAGTACCAAGCTCTCCATTCACAAAGCCAAACG 1729  
DB 421 TGGAGATTCCTTCACTCCACCAAGTACCAAGTACCAAGTACCAAGTACCAAGTACCA 480  
QY 1730 CATCGAGCTTAAGCACTGCTAGTGAAGGGGCGCCCAAGAAAGATCTT 1781  
DB 481 CATCGAGCTTAAGCACTGCTAGTGAAGGGGCGCCCAAGAAAGATCTT 532

RESULT 12  
US-09-339-338-271  
Sequence 271, Application US/09339338A  
Patent No. US2002010262A1  
GENERAL INFORMATION:  
APPLICANT: Yugu, Jiang  
APPLICANT: Dillon, Jennifer L.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
FILE REFERENCE: 210121.470C2  
CURRENT APPLICATION NUMBER: US/09/339,338A  
CURRENT FILING DATE: 1999 06 23  
NUMBER OF SEQ ID NOS: 315  
SOFTWARE: fastseq for Windows Version 3.0  
SEQ ID NO 271  
LENGTH: 533  
TYPE: DNA  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1) (533)  
OTHER INFORMATION: n=A,T,C or G  
US-09-339-338-271

Query Match 11.9%; Score 432.6; DB 10; Length 533;  
Best Local Similarity 88.0%; Pred. No. 5.3e-103;  
Matches 468; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 1250 CCACCGTCACGATATGTCAGCTCACTGCCAAGCGGATGCGCAGGAAAGAACTGCTGG 1309  
DB 1 CCACTGTACGGTCTGTCTGACACTTACTGCCAAAAGCATGCGCAAGAAAAAAGCTGTAG 60  
QY 1310 TGAAGAACCTGGAAGCTGTGGAGACCTTGGGGTCCACATCCACCATCTGCTCCGACAGA 1369  
DB 61 TGAAGAACCTTGAAGAGCTGTGGAGACCTTGGGGTCCACATCCATCTGCTGTATAAA 120  
QY 1370 CTGGAACCTGACTCAGAACCGGATGACAGTGGCTCAGATGGTTGACAAATCAATCC 1429  
DB 121 CTGGAACCTGACTCAGAACCGGATGACAGTGGCTCAGATGGTTGACAAATCAATCC 180  
QY 1430 ATGAAGCTGACACCAAGAGAAATGAGAGTGGAGTCTCTTTGACAAAGCTGACCACT 1489  
DB 181 ATGAAGCTGATACACAGAGAAATGAGAGTGGTGTCTTTTGAACAAGCTTCACTACT 240  
QY 1490 GGTTCGCTCTGTCAGAAATGCTGCTCTCTTAACAGGAGAGTGTTCAGGCTAACCAAG 1549  
DB 241 GGTTCGCTCTGTCAGAAATGCTGCTCTCTTAACAGGAGAGTGTTCAGGCTAACCAAG 300  
QY 1550 AAAACCTGCTATCTTAAAGCTGACAGTGGAGAGATGCTTCCAGTGGGCTCTTAA 1609  
DB 301 AAAACCTGCTATCTTAAAGCTGACAGTGGAGAGATGCTTCCAGTGGGCTCTTAA 360  
QY 1610 AGTCATGAGAGT 1669  
DB 361 AGTCATGAGAGT 420  
QY 1670 TGGAGATTCCTTCACTCCACCAAGTACCAAGCTCTCCATTCACAAAGCCAAACG 1729  
DB 421 TGGAGATTCCTTCACTCCACCAAGTACCAAGTACCAAGTACCAAGTACCAAGTACCA 480  
QY 1730 CATCGAGCTTAAGCACTGCTAGTGAAGGGGCGCCCAAGAAAGATCTT 1781  
DB 481 CATCGAGCTTAAGCACTGCTAGTGAAGGGGCGCCCAAGAAAGATCTT 532

RESULT 13  
US-10-007-805-271  
Sequence 271, Application US/10007805  
Patent No. US2002010581A1  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yugu  
APPLICANT: Dillon, Jennifer L.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Durham, Margarita  
TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
FILE REFERENCE: 210121.470C10  
CURRENT APPLICATION NUMBER: US/10/007,805  
CURRENT FILING DATE: 2001-12-07  
NUMBER OF SEQ ID NOS: 593  
SOFTWARE: fastseq for Windows Version 4.0  
SEQ ID NO 271  
LENGTH: 533  
TYPE: DNA























```

Db 1971 ATACGGCCAAAGGCGCATGGCCAAAGGTTGGGATCATCTCTGAGGCGCAAGAGCTGTG 2030
QY 2161 GAAGAAATTGCTGCTCCGCTCCAAACATTTCAGTGAACGAGGTGAACCCCAAGATGCTAAG 2220
Db 2031 GAGGACATGCGCGCGCGCTCCAAACATTTCAGTGAACGAGGTGAACCCCAAGATGCTAAG 2090
QY 2221 GCTGCTGCTAGTACATGAGCATGAGCTTGAAGAGATGAGCTCTGTGAGAGCTGATGATCATT 2280
Db 2091 GCTGCTGCTAGTACATGAGCATGAGCTTGAAGAGATGAGCTCTGTGAGAGCTGATGATCATT 2150
QY 2281 TTGGCGTACCAACAGAGATGCTGCTTGTGAGAGCTCTGTGAGAGCTGATGATCATT 2340
Db 2151 CTGCAAGATGAGCATGAGCTGCTGCTGAGAGCTCTGTGAGAGCTGATGATCATT 2210
QY 2341 GTGAGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Db 2211 GTGAGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2270
QY 2401 TCTGAGGCTTTGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460
Db 2271 TCTGAGGCTTTGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2330
QY 2461 TCTGAGGCTTTGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
Db 2331 TCTGAGGCTTTGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2390
QY 2521 GTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580
Db 2391 GTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2450
QY 2581 AGTAAATTCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
Db 2451 AGTAAATTCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2510
QY 2641 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700
Db 2511 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2570
QY 2701 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2760
Db 2571 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2630
QY 2761 GAGAAATTCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820
Db 2631 GAGAAATTCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2690
QY 2821 GAGAAATTCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2880
Db 2691 GAGAAATTCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2750
QY 2881 GAGAAATTCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940
Db 2751 GAGAAATTCGGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2810
QY 2941 TACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3000
Db 2811 TACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2870
QY 3001 TTTCTGTCAGTATCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3060
Db 2871 TTTCTGTCAGTATCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2930
QY 3061 AATTCTGCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3120
Db 2931 AATTCTGCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2990
QY 3121 AAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180
Db 2991 AAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3050
QY 3181 CCGCTGAAACCTTACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
Db 3051 CCGCTGAAACCTTACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3110

```

```

QY 3241 GAGGAGGCTGCGGAAGGCTCATCATCAGGCGAAGCGCGCTGAGCGAGCTGTG 3286
Db 3111 GAGGAAATTCGGAAGGCTCATCATCAGGCGAAGCGCGCTGAGCGAGCTGTGAGG 3156

RESULT 2
BC031389
LOCUS BC031389
DEFINITION Mus musculus, clone IMAGE:4946041, mRNA.
ACCESSION BC031389
VERSION BC031389.1
KEYWORDS GI:21595237
SOURCE HTC.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1131)
REFERENCE
AUTHORS Strausberg, P.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-rt@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gumaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Fowles, C.P., Lawrence, S., Martin, P.G., Murny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/BLN at: http://image.llnl.gov
Series: IRAC plate: 60 Row: C Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: incomplete processing.
Location/Qualifiers
1. 1131
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="IMAGE:4946041"
/tissue type="Mammary tumor C3(1)-Tag model. Infiltrating
ductal Carcinoma. 5 month old virgin mouse."
/clone_id="NCI CGAP_Mam6"
/lab_host="DH10B"
/label="Vector: pCMV-Sport6"
BASE COUNT 282 a 280 c 296 g 273 t
ORIGIN

Query Match 29 3%; Score 1027.8; NP 11; Length 1131;
Best Local Similarity 95.7%; Pred. No. 1.9e-261;
Matches 1069; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 2526 AGAAGGCTGTGATTTGATTAAGTAACTGATGCTTACACCTTAACAAGTAA 2585
Db 1 AGAAGGCTGTGATTTGATTAAGTAACTGATGCTTACACCTTAACAAGTAA 60

QY 2586 CATTCGGAATTCACCCCTTGTATTTATTTATTTGCAACATTCACGCGCCCTGGG 2645
Db 61 CATTCGGAATTCACCCCTTGTATTTATTTATTTGCAACATTCACGCGCCCTGGG 120

QY 2646 CACGCTGACCACTGCTGTGATTTGATTTGAGGAGCTGACATGATTTCCGCGGATCTGTGAGC 2705

```



```

Db 121 GACTGTGACGACCTCTGATGACTTGGGCACTGACATGTTCCCTGCGCTTCCTTGGC 180
QY 2706 CTATGAAAGG3TGAAGTGAATATGAAAGAGGCGGCTTAAATAATCCAAAAG3ACAA 2765
Db 181 CTACAGGAGAGGCTGAGAGGAGCATCATGAAAGAGGAGCCGAGAAACCCGAAAAC3ACAA 240
QY 2766 ACTGTGAAGGAGGCTGATGAGCATGAGGAGGAGTGAAGAGATGATGATGAGGAGGCT 2825
Db 241 ATTGTGAAGGAGGCTGATGAGCATGAGGAGGAGTGAAGAGATGATGATGAGGAGGCT 300
QY 2826 GAGAGGCTGATGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2885
Db 301 GAGAGGCTGATGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 2886 GTTGGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2945
Db 361 GTTGGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 2946 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3005
Db 421 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 3006 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3065
Db 481 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 3066 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3125
Db 541 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 3126 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3185
Db 601 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 3186 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3245
Db 661 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 3246 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3305
Db 721 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 3306 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3365
Db 781 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 3366 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3425
Db 841 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 3426 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3479
Db 901 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 3480 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3539
Db 961 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 3540 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3599
Db 1021 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 3600 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3636
Db 1081 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1117

```

```

RESULT 3
LOCUS BC023175 1009 bp mRNA linear HTC 07-AUG 2002
DEFINITION Mus musculus, clone IMAGE:4952025, mRNA.
ACCESSION BC023175

```

```

VERSION BC023175.1 GI:18606022
KEYWORDS HTC
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 1009)
AUTHORS Strausberg, P.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgabs@imail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLUM)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.P., Lawrence, S., Martin, P.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLUM at: <http://image.lln.gov>  
 Series: IRAC Plate: 60 Row: 1 Column: 21  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis  
 This clone has the following problem: incomplete processing

## FEATURES

```

source
  1..1009
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /map="FVB/N"
    /clone="IMAGE:4952025"
    /tissue_type="Mammary tumor, C3(1)-Tag model Infiltrating
    ductal carcinoma, 5 month old virgin mouse."
    /clone_1b="NCI CGAP Mam6"
    /lab_host="DH10B"
    /note="Vector: PCMV-SPORT6"
BASE COUNT      250 a      247 c      276 g      236 t
ORIGIN

```

Query Match 25.0%; Score 910.8; DB 11; Length 1009;

Best Local Similarity 95.7%; Pred. No. 2.2e-230; Mismatches 37; Indels 6; Gaps 1;

```

Db 2851 TGACATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2710
QY 1 TGACATCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
Db 2711 AACAGCTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2770
QY 2711 AACAGCTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2770
Db 61 AGCAAGCTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 2771 TGAAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2830
Db 121 TGAAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 2831 GCTTCTTCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2890
Db 181 GCTTCTTCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 2891 GCATCCGAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2950
Db 241 GCATCCGAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300

```



QY	2951	AGTGGACCTTACGAGCAGAGGAGAAAGATTGTGGAGTTCACTGTCACACGAGCCCTCTTTGTCA	3010
Db	301	AGTGGACCTTACGAGCAGAGGAGAAAGATTGTGGAGTTCACTGTCACACGAGCCCTCTTTGTCA	360
QY	3011	GTATCGTGTAGTGCACATGGGCTGACTTGGTTCATCTTCCAAAGCCAGAAAGAAATTCTGTCT	3070
Db	361	GTATTGTGTAGTGCACATGGGCTGACTTGGTTCATCTTCCAAAGCCAGAAAGAAATTCTGTCT	420
QY	3071	TCCAGCAGGGAAATGAAAGAAACAATCTTTAATATTTGGCCCTCTTTGAAGAAGACAGCTCTTG	3130
Db	421	TCCAGCAGGGAAATGAAAGAAACAATCTTTAATATTTGGCCCTCTTTGAAGAAGACAGCTCTTG	480
QY	3131	CTGCTTCCGCTGCTCTTACTGCGCCCTGGATGGGTCAGCCCTTAGATGTATCCCTCAAC	3190
Db	481	CTGCTTCCGCTGCTCTTACTGCGCCCTGGATGGGTCAGCCCTTAGATGTATCCCTCAAC	540
QY	3191	CTACTTGGTGGTCTGTGCTCTTCCCTTACTCTCTTCTCATCTTGTGTATGAGAGGTGC	3250
Db	541	CTACTTGGTGGTCTGTGCTCTTCCCTTACTCTCTTCTCATCTTGTGTATGAGAGGTGC	600
QY	3251	GGAAAGCTCATCATCAGAGGAGACGCGCTGGCTGGGTGGAGAGAGAAACCTTACTAGC	3310
Db	601	GGAAAGCTCATCATCAGAGGAGACGCGCTGGCTGGGTGGAGAGAGAAACCTTACTAGC	660
QY	3311	CCACTGCGCTGACAGCGCGTGGACATTTGTGCCACACACTGACCTTACCCTTACCCCTCT	3370
Db	661	CCACTGCGCTGACAGCGCGTGGACATTTGTGCCACACACTGACCTTACCCTTACCCCTCT	720
QY	3371	TTTGTTACTTCAAGCTCTTTGGAGCTGGAAACCTTACCCCTGTAGAGAAACACCAAGCATG	3430
Db	721	TTTGTTACTTCAAGCTCTTTGGAGCTGGAAACCTTACCCCTGTAGAGAAACACCAAGCATG	780
QY	3431	TGGGATCTCAGACGCTCTGGAATGAAAGCATGTAGCTGTATGGGAGGCG-----GGGAGG	3484
Db	781	TGGGATCTCAGACGCTCTGGAATGAAAGCATGTAGCTGTATGGGAGGCGGGGAGGAGG	840
QY	3485	AGGGCTGCGCCGAAAAACCGCTGGACGGGGACGACAGCGGGGAAGCTTTATATGGCCCT	3544
Db	841	AGGGCTGCGCCGAAAAACCGCTGGACGGGGACGACAGCGGGGAAGCTTTATATGGCCCT	900
QY	3545	TTTGTCTTTGTAAAAAGAAAAACCTGGAAAAAGACTGAAAGATTACGTTTATATCTGGAT	3604
Db	901	TTTGTCTTTGTAAAAAGAAAAACCTGGAAAAAGACTGAAAGATTACGTTTATATCTGGAT	960
QY	3605	TTTTACAAATAAAGATGGCTATTATATACGGAA	3636
Db	961	TTTTACAAATAAAGATGGCTATTATATACGGAA	992
RESULT 4			
LOCUS	BQ715991	935 bp	mRNA
DEFINITION	AGNCOURT_8294958 NIH_MGC_129	Mus musculus	CDNA clone IMAGE:6309731
ACCESSION	BQ715991		
VERSION	BQ715991.1	GI:21854890	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Mumaiyola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Emmaliya; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 935)		
JOURNAL	NIH-MGC http://mgi.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Susan L. Sullivan, PhD.		
	cDNA Library Preparation: Resgen, Invitrogen Corp		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		

FEATURES									
source									
1. .935									
/organism="Mus musculus"									
/db_xref="taxon:10090"									
/clone="IMAGE:6309731"									
/clone_1id="NIH MGC 129"									
/lab_host="DH10B (phage-resistant)"									
/note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1.cdb, Site_1: EORV; Site_2: NotI; Cloned unidirectionally. Primer: O130 dt. Average insert size 2.2 kb. Constructed by Resgen. In vitro library. Note: This is a NIH MGC Library."									
BASE COUNT 227 a 240 c 251 g 216 t 1 others									
ORIGIN									
Query Match 21.6%; Score 786.8; DB 14; Length 935;									
Best Local Similarity 93.3%; Pred. No. 1.8e-197;									
Matches 854; Conservative 0; Mismatches 58; Indels 3; Gaps 3									
QY	2183	ACATTCGAGTACACGAGGTGAACCCACAGATGCGACAGCCCTGTGTATACATGACATG	2242						
Db	1	ACATTCGAGTACACGAGGTGAACCCACAGATGCGACAGCCCTGTGTATACATGACATG	60						
QY	2243	ACTTGAAAGACATGACCTCTTATGAGAGCTTGATGATATTTTGGGTTACACAGAGATTTG	2302						
Db	61	ACTTGAAAGACATGACCTCTTATGAGAGCTTGATGATATTTTGGGTTACACAGAGATTTG	120						
QY	2303	TCTTTGCTAGGACCTCTCTCTCAACAGAGCTCATATTGTGGAGGGCTGCCAGCGGACAG	2362						
Db	121	TCTTTGCTAGGACCTCTCTCTCAACAGAGCTCATATTGTGGAGGGCTGCCAGCGGACAG	180						
QY	2363	GTGCCATCGTGGCTGTCAAGGGAGTGTGTCATGACTCTCCAGCTTTGAAAAAGCGAG	2422						
Db	181	GTGCCATCGTGGCTGTCAAGGGAGTGTGTCATGACTCTCCAGCTTTGAAAAAGCGAG	240						
QY	2423	ATATTGGGGTTGCCATGGGGATTTGTGGTGGATGTGTCCAAAGCAAGTGGTGCATCATGA	2482						
Db	241	ATATTGGGGTTGCCATGGGGATTTGTGGTGGATGTGTCCAAAGCAAGTGGTGCATCATGA	300						
QY	2483	TTCTCTCGATACACAATCTTGGCTCCATCGTACTGAGATAGAGAGGTGCTCTGATAT	2542						
Db	301	TTCTCTCGATACACAATCTTGGCTCCATCGTACTGAGATAGAGAGGTGCTCTGATAT	360						
QY	2543	TTGATPACTTGAAGAAATCATTTGCTTACACCTTACAAAGTAACATTCGGGAATATACCC	2602						
Db	361	TTGATPACTTGAAGAAATCATTTGCTTACACCTTACAAAGTAACATTCGGGAATATACCC	420						
QY	2603	CCTCTCTGATATTTATTTATTTGCAAAATTCACACTGCCCCCTGGGACACCGTACATCCTCT	2662						
Db	421	CCTCTCTGATATTTATTTATTTGCAAAATTCACACTGCCCCCTGGGACATGACATCCTCT	480						
QY	2663	GCATGACTTGGGCACTGACATGTTTCCGCCATCTCTTGCGCTTATGAACAGGCTGAAA	2722						
Db	481	GCATGACTTGGGCACTGACATGTTTCCGCCATCTCTTGCGCTTATGAACAGGCTGAGA	540						
QY	2723	GTGACATCATGAAAGGGGCGAGCCAGAAATCCCAAAACGCAAACTTGTGAACGAGCGTC	2782						
Db	541	GTGACATCATGAAAGGGGCGAGCCAGAAATCCCAAAACGCAAACTTGTGAACGAGCGTC	600						
QY	2783	TGATCGACATGGGCTATGAGACAGATGGATGATGATCCAGGCGCTTGGAGAGCTCTTCACTT	2842						
Db	601	TGATCGACATGGGCTATGAGACAGATGGATGATGATCCAGGCGCTTGGAGAGCTCTTCACTT	660						
QY	2843	ATTTTGTGATTCGTGGCTGAGAAAGGTTTCTGCGCTTTACCTGTGTGGGATCCGAGAGA	2902						
Db	661	ATTTTGTGATTCGTGGCTGAGAAAGGTTTCTGCGCTTTACCTGTGTGGGATCCGAGAGA	720						
QY	2903	CTGGGAGTACCGCTGATTAATGATGTGGAGACAGCTACGGGACAGCACTGACCTACG	2962						
Db	721	CTGGGAGTACCGCTGATTAATGATGTGGAGACAGCTACGGGACAGCACTGACCTACG	780						







CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM13744 row: 1 column: 09  
 High quality sequence stop: 607.

## FEATURES

Location/Qualifiers

1..916  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5256937"  
 /lab\_host="NIH MGC 129"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: olfactory epithelium; Vector:  
 pCMV-SPORT6.1.cdb; Site\_1: EcoRI; Site\_2: NotI; Cloned  
 unidirectionally. Primer: Oligo dt. Average insert size  
 2.2 kb. Constructed by Resgen, Invitrogen Corp. Note: this  
 is a NIH MGC Library."  
 BASE COUNT 209 a 252 c 243 g 211 t 1 others  
 ORIGIN  
 Query Match 21.1%; Score 767.4; DB 14; Length 916;  
 Best Local Similarity 92.7%; Pred. No. 2.6e-192;  
 Matches 849; Conservative 0; Mismatches 62; Indels 5, Gaps 4;

1756 ATGAGAGGCGCCCGAGAAAGATCTGACCGATGAGCTTATCTCCCTCCAGCGGCAAG 1815  
 1 ATGAGAGGCGCCCGAGAAAGATCTGACCGATGAGCTTATCTCCCTCCAGCGGCAAG 60  
 1816 GACGAGCGCCCTGAGCAAGAGCTGAGAGAGCGCTTTCAGAAATGCTTACCTAGAGCTGGG 1875  
 61 GAGAGGCACTGATGTAAGAGCTGAGAGAGCGCTTTTGAAGCGCTTACCTAGAGCTGGG 120  
 1876 GGGCTTGAAGAGCTGCTGCTGCTTCTGCACTCCTCTCTGCTGAGCAAGATTTCC 1935  
 121 GGGCTTGAAGAGCTGCTGCTGCTTCTGCACTCCTCTCTGCTGAGCAAGATTTCC 180  
 1936 GAAAGGCTTCCAGTTGACATGATGAGATGATTTCCCGTGAGATTAACCTGCTGCTG 1995  
 181 GAAAGGCTTCCAGTTGATGATGAGATGATTTCCCGTGAGATTAACCTGCTGCTG 240  
 1996 GGTCTTATCTGATGATGATGATGATTTCCCGTGAGATTTCCCGTGAGATTTCC 2055  
 241 GGTCTTATCTGATGATGATGATTTCCCGTGAGATTTCCCGTGAGATTTCC 300  
 2056 GCGAGCGCTGGGATTAAGGATGATGATGATTTCCCGTGAGATTTCCCGTGAGATTTCC 2115  
 301 GCGAGCGCTGGGATTAAGGATGATGATGATTTCCCGTGAGATTTCCCGTGAGATTTCC 360  
 2116 ATTGCTAAGGAGGAGTATGATGATGATTTCCCGTGAGATTTCCCGTGAGATTTCC 2175  
 361 ATTGCTAAGGAGGAGTATGATGATGATTTCCCGTGAGATTTCCCGTGAGATTTCC 420  
 2176 CGGCTCAACATTTCCAGTAAACCGAGTAAACCGAGATTTCCCGTGAGATTTCC 2235  
 421 CGGCTCAACATTTCCAGTAAACCGAGTAAACCGAGATTTCCCGTGAGATTTCC 480  
 2236 GCGAGTACTGGAAGATGATGATGATTTCCCGTGAGATTTCCCGTGAGATTTCC 2295  
 481 GCGAGTACTGGAAGATGATGATTTCCCGTGAGATTTCCCGTGAGATTTCC 540  
 2296 GAGATTTGCTTTGCTAGAGCTCTCTCAAGAGAGATTTGTTGAGAGGCTGGCAG 2355  
 541 GAGATTTGCTTTGCTAGAGCTCTCTCAAGAGAGATTTGTTGAGAGGCTGGCAG 600  
 2356 GAGATTTGCTTTGCTAGAGCTCTCTCAAGAGAGATTTGTTGAGAGGCTGGCAG 2414  
 601 GAGATTTGCTTTGCTAGAGCTCTCTCAAGAGAGATTTGTTGAGAGGCTGGCAG 660  
 2415 AAAGGCAATATTTGAGGATTTGCTAGAGGATTTGCTAGAGGATTTGCTAGAGGATTTG 2474  
 661 AAAGGCAATATTTGAGGATTTGCTAGAGGATTTGCTAGAGGATTTGCTAGAGGATTTG 720

QY 2475 TGACATGATTTCTTGATGATGACAACTTTGCTCCATGCTGATGGA-GTGAAGAAGGTC 2533  
 DB 721 TGACATGATTTCTTGATGATGACAACTTTGCTCCATGCTGATGGA-GTGAAGAAGGTC 780  
 QY 2534 GTCGATATTTGATTAACCTTGAAGAAATCAATGCTTACACCTTCAAGATTAACATTTCCG 2593  
 DB 781 GTCGATATTTGATTAACCTTGAAGAAAT-CAATGCTTACACCTTCAAGATTAACATTTCCG 839  
 QY 2594 AAATACCCCTCTTGAATATTTATTTGCAAAATTTCACTGCCCC--TGCGACCCGT 2651  
 DB 840 AAATACCCCTCTTGAATATTTATTTGCAAAATTTCACTGCCCCCTGCGGACTGTG 899  
 QY 2652 GACCATCTCTGCAAT 2667  
 DB 900 AACCATCTCTGCAAT 915

RESULT 7  
 BI904878 895 bp mRNA linear EST 16-OCT-2001  
 LOCUS 603168903F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5256937 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI904878  
 VERSION BI904878.1 GI:16167292  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: gcapbs@mail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM1649 row: e column: 02  
 High quality sequence start: 58  
 High quality sequence stop: 893.  
 Location/Qualifiers

1..895  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5256937"  
 /lab\_host="NIH CGAP Lu33"  
 /tissue\_type="pooled lung tumors"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: NotI; Site\_2: EcoRI; 1st  
 strand cDNA was prepared from mRNA obtained from pooled  
 lung tumors with a Not I - oligo(dT) primer [5',  
 TGTACCAATCTGAAAGTGGAGCGCGCGCTCTGTTTTTTTTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 211 a 237 c 245 g 232 t  
 ORIGIN

Query Match 20.7%; Score 753.6; DB 13; Length 895;  
 Best Local Similarity 94.0%; Pred. No. 1.2e-188;  
 Matches 805; Conservative 0; Mismatches 49; Indels 2, Gaps 2;







```
Query Match      20.7%  Score 751.8;  DB 14;  Length 947;
Best Local Similarity  91.1%  Pred. No. 3.7e+188;
Matches  856;  Conservative  0;  Mismatches  70;  Indels  14;  Gaps  5;
```

[illegible]















RESULT 14  
BO889413  
LOCUS  
DEFINITION  
AGENCOURT\_8728430 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:6339047  
5', mRNA sequence.  
ACCESSION  
BO889413  
VERSION  
BO889413.1 GI:22281427  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens  
ORGANISM  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
1 (bases 1 to 933)  
NIH-MGC <http://imgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LNCM253 row: d column: 24  
High quality sequence stop: 691.  
Location/Qualifiers  
1..933  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6339047"  
/clone\_id="NIH\_MGC\_47"  
/tissue\_type="neuroblastoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: brain; Vector: pOT87; Site: 1: XhoI; Site: 2:  
EcoRI; cDNA made by oligo-dt priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH-MGC Library."  
BASE COUNT 239 a 217 c 234 g 239 t 4 others  
ORIGIN

Query Match 20.2%; Score 735.4; DB 14; Length 933;  
Best Local Similarity 88.8%; Pred. No. 8.3e-184;  
Matches 793; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 717 CTTCAGAGATGCTCCTCAGCAAGCCTCGTATTCGAATGAGAGAGATGAGCAT 776  
Db 4 CTTCAAAAGATGCTCCTCAGCAAGCCTCGTATTCGAATGAGAGATGAGCAT 63  
QY 777 CAAGCAGAGATGCTCCTGTTGATCTGTGAGGTGAAGGCGAGACCGAATCC 836  
Db 64 AAATGCGAGAGATGTTGTTGGGAGATCTGTGGAAGTAAGAGAGACCGAATCC 123  
QY 837 TGGTGAATTCAGATTCATATTCGAAATGAGGTGCAAGTGAATCTCTCTCACTAGG 896  
Db 124 TGCTGATTCAGATTCATATTCGAAATGAGGTGCAAGTGAATCTCTCTCACTAGG 183  
QY 897 TGAATGAGAACCCGAGACTGGTCCCGGATTTTCAGAAAGAGAACCCCTGGAGACAAG 956  
Db 184 TGAATGAGAACCCGAGACTGGTCCCGGATTTTCAGAAATGAGAACCCCTGGAGACAAG 243  
QY 957 GAACATTTGCTTTCTTCAACCAACTGTGTGAAGAACTGACGCTGAGATGTTGTGA 1016  
Db 244 GAACATTTGCTTTCTTCAACCAACTGTGTGAAGAACTGACGCTGAGATGTTGTGA 303  
QY 1017 CACTGGGATTCGACCTGTGATGGGAGAGATGCCACCTTGTGGCTGGAAGGCGG 1076

Db 304 CACTGGGATTCGACCTGTGATGGGAGAGATTCGACACTTGTCTTGGGCTGAAGAGAG 363  
QY 1077 CCGACCCCATTTGCTGTAAGAAATCGAGCACTTATCACCCTCTTCAGGCTGTCCTG 1136  
Db 364 CGAGACCCCATTTGCTGTAAGAAATCGAGCACTTATCACCCTCTTCAGGCTGTCCTG 423  
QY 1137 GTTCCGAGGAGTGTCTTTCTTCAATTCCTCTGTGATCCTTAAATACACCTGAGTGAAG 1196  
Db 424 GTTCCGAGGAGTGTCTTTCTTCAATTCCTCTGTGATCCTTAAATACACCTGAGTGAAG 483  
QY 1197 TGTGATCTTCTCATTTGATGATTCATGTCAGCCACGTCGCGGAGAGTTTCTGCGCCACGT 1256  
Db 484 TGTGATCTTCTCATTTGATGATTCATGTCAGCCACGTCGCGGAGAGTTTCTGCGCCACGT 543  
QY 1257 CAGGATATGCTGACGCTGCTGTCGCAAGGCTGTCGCAAGGCTGTCGCAAGGCTGTCG 1316  
Db 544 CAGGATATGCTGACGCTGCTGTCGCAAGGCTGTCGCAAGGCTGTCGCAAGGCTGTCG 603  
QY 1317 CCTGGAAGCTGTGAGACCTTGGGGTCCACATCCACATCTGCTCCGACAGACTGGAAC 1376  
Db 604 CTTAAGAGCTGTGAGACCTTGGGGTCCACATCCACATCTGCTCCGACAGACTGGAAC 663  
QY 1377 TCTGACTCAAGCCGATGATGACAGTGTGCTCAATGTGTTTACATCAATTCATGAAGC 1436  
Db 664 TCTGACTCAAGCCGATGATGACAGTGTGCTCAATGTGTTTACATCAATTCATGAAGC 723  
QY 1437 TGACACACAGAGATGAGTGGGGTCTCCTTGAACAAGCGTCAGCTGCTGCTGGC 1496  
Db 724 TGATGACACAGAGATGAGTGGGGTCTCCTTGAACAAGCGTCAGCTGCTGCTGGC 783  
QY 1497 TCTGTCAGATTTGCTGCTCTGTGTAACAGGCGAGTGTTCAGGCTAACAAGAAAGCT 1556  
Db 784 TCTGTCAGATTTGCTGCTCTGTGTAACAGGCGAGTGTTCAGGCTAACAAGAAAGCT 843  
QY 1557 GCTATTCCTTAAGCGTGCAGTAGCGGAGATGCTTCGAGTGGCGCTTTAA 1609  
Db 844 ACCTATTTCTTAAGCGTGCAGTAGCGGAGATGCTTCGAGTGGCGCTTTAA 996

RESULT 15  
BM557003  
LOCUS  
DEFINITION  
AGENCOURT\_6578791 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5465507  
5', mRNA sequence.  
ACCESSION  
BM557003  
VERSION  
BM557003.1 GI:18798681  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
1 (bases 1 to 1032)  
NIH-MGC <http://imgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: DCTD/DRP  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LNCM1967 row: a column: 04  
High quality sequence stop: 726.  
Location/Qualifiers  
1..1032  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5465507"  
/clone\_id="NIH\_MGC\_41"  
/tissue\_type="amelanotic melanoma, cell line"



